

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2005, 08:52:10 ; Search time 291 Seconds

(without alignments)  
2925.772 Million cell updates/sec

Title: US-10-615-144-3

Perfect score: 1737

Sequence: 1 gaattcgccgcgcattga.....ataatcgccgcgcgaattc 1737

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4161359 seqs, 245077644 residues

Total number of hits satisfying chosen parameters: 8322718

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA New:\*

- 1: /cgn2\_6/ptodata/1/pubna/US09\_NEW\_PUB.seq.\*
- 2: /cgn2\_6/ptodata/1/pubna/US06\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/1/pubna/US07\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/1/pubna/US08\_NEW\_PUB.seq.\*
- 5: /cgn2\_6/ptodata/1/pubna/PCT\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/1/pubna/US10\_NEW\_PUB.seq.\*
- 7: /cgn2\_6/ptodata/1/pubna/US11\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/1/pubna/US11\_NEW\_PUB.seq.\*
- 9: /cgn2\_6/ptodata/1/pubna/US11\_NEW\_PUB.seq.\*
- 10: /cgn2\_6/ptodata/1/pubna/US60\_NEW\_PUB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	146.4	8.4	3109	US-10-821-234-63	Sequence 63, Appl
C 2	41.6	2.4	977	US-10-750-185-58338	Sequence 58338, A
C 3	36.6	2.1	2725	US-10-131-826A-479	Sequence 479, App
C 4	35.4	2.0	1082144	US-11-117-187-211	Sequence 211, App
C 5	34.6	2.0	215308	US-11-121-086-77	Sequence 77, Appl
C 6	34.2	2.0	747	US-10-750-185-57746	Sequence 57746, A
C 7	34.2	2.0	80450	US-11-117-187-201	Sequence 201, App
C 8	34	2.0	1276	US-10-750-185-33227	Sequence 33227, A
C 9	34	2.0	3619	US-10-793-626-4157	Sequence 4157, Ap
C 10	34	2.0	645179	US-10-995-561-13293	Sequence 13293, A
C 11	33.8	1.9	1056	US-10-750-185-52662	Sequence 52662, A
C 12	33.8	1.9	2460	US-10-793-626-4439	Sequence 4439, Ap
C 13	33.8	1.9	2683	US-10-793-626-3709	Sequence 3709, Ap
C 14	33.8	1.9	3022	US-10-793-626-3969	Sequence 3969, Ap
C 15	33.8	1.9	4039	US-10-793-626-3361	Sequence 3361, Ap
C 16	33.6	1.9	794	US-10-750-185-52150	Sequence 52150, A
C 17	33.6	1.9	1432	US-10-750-185-56595	Sequence 56595, A
C 18	33.6	1.9	14341	US-10-995-561-13428	Sequence 13428, A
C 19	33.6	1.9	398287	US-10-995-561-13396	Sequence 13396, A
C 20	33.4	1.9	685	US-10-750-185-34681	Sequence 34681, A
C 21	33.4	1.9	1936	US-11-054-168B-15	Sequence 15, Appl
C 22	33.2	1.9	150437	US-11-112-908-44	Sequence 44, Appl
C 23	33.2	1.9	182314	US-11-112-908-45	Sequence 45, Appl

C 24	33	1.9	1198	6	US-10-750-185-54391	Sequence 54391, A
C 25	33	1.9	1437	6	US-10-793-626-767	Sequence 767, App
C 26	33	1.9	1650	6	US-10-750-185-39856	Sequence 39856, A
C 27	33	1.9	2129	6	US-10-750-185-40823	Sequence 40823, A
C 28	33	1.9	2171	6	US-10-750-185-28678	Sequence 28678, A
C 29	33	1.9	3710	6	US-10-793-626-3404	Sequence 3404, Ap
C 30	33	1.9	31028	6	US-10-829-826B-21	Sequence 21, Appl
C 31	33	1.9	31028	6	US-10-829-826B-22	Sequence 22, Appl
C 32	33	1.9	31028	6	US-10-829-826B-26	Sequence 26, Appl
C 33	33	1.9	31032	6	US-10-829-826B-23	Sequence 23, Appl
C 34	33	1.9	31100	6	US-10-829-826B-24	Sequence 24, Appl
C 35	32.8	1.9	1695	6	US-10-750-185-40850	Sequence 40850, A
C 36	32.8	1.9	159695	7	US-11-121-086-56	Sequence 56, Appl
C 37	32.6	1.9	207600	7	US-11-112-908-31	Sequence 31, Appl
C 38	32.4	1.9	959	6	US-10-750-185-36299	Sequence 36299, A
C 39	32.4	1.9	1726	6	US-10-750-185-50808	Sequence 50808, A
C 40	32.4	1.9	3394	6	US-10-793-626-3365	Sequence 3365, Ap
C 41	32.4	1.9	94618	7	US-11-117-187-191	Sequence 191, App
C 42	32.4	1.9	96128	6	US-10-995-561-13197	Sequence 13197, A
C 43	32.2	1.9	948	6	US-10-750-185-35671	Sequence 35671, A
C 44	32.2	1.9	991	6	US-10-750-185-48683	Sequence 48683, A
C 45	32.2	1.9	1910	6	US-10-750-185-37618	Sequence 37618, A

ALIGNMENTS

RESULT 1

US-10-821-234-63/c  
; Sequence 63, Application US/10821234  
; Publication No. US20050255114A1  
; GENERAL INFORMATION:  
; APPLICANT: Labat, Ivan  
; APPLICANT: Stache-Crain, Birgit  
; APPLICANT: Andarmani, Susan  
; APPLICANT: Tang, Y. Tom  
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Preeclampsia  
; CURRENT APPLICATION NUMBER: US/10/821,234  
; CURRENT FILING DATE: 2004-04-07  
; PRIOR APPLICATION NUMBER: US 60/462,047  
; PRIOR FILING DATE: 2003-04-07  
; NUMBER OF SEQ ID NOS: 1704  
; SOFTWARE: pt\_seq\_genes Version 1.0  
; SEQ ID NO 63  
; LENGTH: 3109  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-821-234-63

Query Match	8.4%	Score 146.4;	DB 6;	Length 3109;
Best Local Similarity	51.1%	Pred. No. 7.3e-34;		
Matches	373;	Conservative	0;	Mismatches 351;
				Indels 6;
				Gaps 1;
QY	676	TACTACAAAATTGCACGTCATTACAAAGTGGGCAATGATGATCAGCTGTTTACAAAGATAAT	735	
Db	1893	TACTACAAAGATCGCGCCACTACCGCTGGCGCTGGCGCAGGTCTTCGGCAGTTTCGC	1834	
QY	736	TTTAGCGGTGTTATCATACTAGAACATGATGAAATTCGCCCTGATTTTTCGACTTT	795	
Db	1833	TTCCCGCGCGCGCTGGTGGTGGAGATGACCTGGAGGTGGCCCTCGGACTTCTTCGAGTAC	1774	
QY	796	TTTAGAGGTGGAGCTACTCTTCTTCACAGACAGACAGTTCGATATGCTATTTCTTTGG	855	
Db	1773	TTTCGGGCAACCTATCCGCTGCTGAAGCCGACCCCTCCCTGTGGTGGCTCGGCTGG	1714	
QY	856	AATGACAAATGGAATAATGAGTTTGTCCAAAG-----ATCCTTATGCTCTTTACCGCTCA	909	
Db	1713	AATGACACGGCAAGAGCAGATGTTGGAGCGCCAGCAGCTGAGCTGCTCTACCGCACC	1654	
QY	910	GATTTTTTCCCGGCTTGGATGGATGCTTTCAAATCTACTTTGGGACGAATATCTCCA	969	
Db	1653	GACTTTTTCCCTGGCTGGCTGTGTGTGGCGAGCTCTGGCTGAGCTGGAGCCC	1594	



Db 1293 TGCTGTGGTCTTGAAGAGGACCTGGACATGCTGTGGATTTTTCAGTTTCTCTGAGCCA 1352

Qy 804 TGGAGCTACTCTTCTTGAAGAGGACCTGGACATGCTGTGGATTTTTCAGTTTCTCTGAGCCA 863

Db 1353 ATCCATCCACTACTGGAGGAGTACAGCCTGTACTGCTCTCTGCTGGAATGACCA 1412

Qy 864 TGGACAAATGAGTTTCTCAAGATCTTATGCTCTTTTACGCTTCAGATTTTTCCTCGG 923

Db 1413 GGGGTATGAACACAGCTGAGGCCAGCAGACTACTGCTGCTGAGAGCAATGCTCTGG 1472

Qy 924 TCTTGGATGGATGCT 938

Db 1473 GCTGGCTGGTGTCT 1487

RESULT 4

US-11-117-187-211/c

Sequence 211, Application US/11117187

Publication No. US20050266560A1

GENERAL INFORMATION:

APPLICANT: PREUSS, DAPHNE

APPLICANT: COPELHAY, GREGORY

TITLE OF INVENTION: PLANT ARTIFICIAL CHROMOSOME COMPOSITIONS AND METHODS

FILE REFERENCE: ARCD:309US

CURRENT APPLICATION NUMBER: US/11/117,187

CURRENT FILING DATE: 2005-04-28

PRIOR APPLICATION NUMBER: US/09/531,120

PRIOR FILING DATE: 2000-03-17

PRIOR APPLICATION NUMBER: 60/125,219

PRIOR FILING DATE: 1999-03-18

NUMBER OF SEQ ID NOS: 212

SOFTWARE: Patent In Ver. 2.1

SEQ ID NO 211

LENGTH: 1082144

TYPE: DNA

ORGANISM: Arabidopsis thaliana

US-11-117-187-211

Query Match 2.0%; Score 35.4; DB 7; Length 1082144;

Best Local Similarity 45.4%; Pred. No. 1e+02;

Matches 164; Conservative 0; Mismatches 196; Indels 1; Gaps 1;

Qy 498 TATTAAATCCATCTTAAATACCAATATCTGTTGGTCAAAATATCTCTTTTCATATC 557

Db 497648 TAAATATTAGGTTAGAAATAACTAATATTTTAAATATATATTTTGGATTTT 497589

Qy 558 CCAGATGGATCACAATCTGATGTGAGGAAGCTTCTTTGAGCTATGATCAGCTGACGTA 617

Db 497588 CGGTATCCAAATCTTTTGGTCTCTATCCATTCGGTTTATTTAAATTTATGTT 497529

Qy 618 TATGAGCACTTGGATTTTGAACCTGTGCTACTGAAAGACAGGGAGCTGATTT-GCAT 676

Db 497528 TCGGTTTCAGTTCGGATTCGGTTTAAAAATTTTAAAAACCGCATCGAATTCGGTT 497469

Qy 677 ACTACAAATTCACGCTATTACAGTGGGATTCGATCAGCTGTTTACAGCAATAAT 736

Db 497468 ATTTTAAATTTATGTTTCGGTTTCAGTTCGGATTCGGTTTAAATTTATATCTATT 497409

Qy 737 TTAGCGGTGTTATCATCTACTAGAGATGATATGAAATTCGCCCTGATTTTTCATTTT 796

Db 497408 TTATTTTGTGTTAAATATTTTGTAGATATTTTGTAGTTATTTTGTATATTTTTCATATTTT 497349

Qy 797 TTGAGGCTGGAGCTACTCTTTTGAAGAGCAAGTTCGATTTATGCTATTTCTTCTTGA 856

Db 497348 GGTACTTTCGAGTATTTTACTATTTATACGTATATCTATTTTAAATTTATAGTTTGA 497289

Qy 857 A 857

Db 497288 A 497288

RESULT 5

US-11-121-086-77/c

Query Match 2.0%; Score 34.2; DB 6; Length 747;

Best Local Similarity 55.5%; Pred. No. 3.1;

Matches 66; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

Qy 1275 TGTGCTATTTCAGTACAGAGATCACTAGACTTTTGAATAATTCGACCGCAATTTGGCAT 1334

Db 368 TTTATGTTTGAACACAGACACTAAATGACATTTTATATGTCGATCGACAGTTTGTCT 309

Qy 1335 TTTTGAAGAATGGAAGGATGTTACCGCTGACATATTAAGGAATAGTAGTTTTC 1393

Db 308 TTTATTTGTGAATATGATTAATCTTGTCTTTTCAAGGGTTTAGGTAGACTGGCTTTC 250

Sequence 77, Application US/11121086

Publication No. US20050266459A1

GENERAL INFORMATION:

APPLICANT: POULSEN, TIM S.

APPLICANT: NIELSEN, KIRSTEN V.

TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES

FILE REFERENCE: 09138.6000-00000

CURRENT APPLICATION NUMBER: US/11/121,086

CURRENT FILING DATE: 2005-05-04

PRIOR APPLICATION NUMBER: 60/567,570

PRIOR FILING DATE: 2004-05-04

NUMBER OF SEQ ID NOS: 107

SOFTWARE: Patent In version 3.3

SEQ ID NO 77

LENGTH: 215308

TYPE: DNA

ORGANISM: Homo sapiens

US-11-121-086-77

Query Match 2.0%; Score 34.6; DB 7; Length 215308;

Best Local Similarity 52.4%; Pred. No. 75;

Matches 76; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

Qy 1017 TCACAGAGGTGCAAAATTTATTTCCCCAGAGTTTGCAGAACATATATTTTGGTACCA 1076

Db 91543 TTACAATTTTGTCCCTTCTCTGTTTCTAAATTTTGTAAATATATACCATTTTCTCTGATTT 91484

Qy 1077 TGGTCTAGTTTGGGCGAGTTTTCAGAGCATATCTTTGAGCAATTAACATAATGATGT 1136

Db 91483 CCTTAGTTTTCCTCCATGTTTCTTTTACTCTTTGAGCAAAATTTAAGCAATTTGTTT 91424

Qy 1137 CCAGTTTGATGGAAGTCAATGGAC 1161

Db 91423 TAAAGTCTTGTAAAGTCAATGTC 91399

RESULT 6

US-10-750-185-57746/c

Sequence 57746, Application US/10750185

Publication No. US20050260603A1

GENERAL INFORMATION:

APPLICANT: MMI GENOMICS, INC.

APPLICANT: DENISE, Sue K.

APPLICANT: KERR, Richard

APPLICANT: ROSENFELD, David

APPLICANT: HOLM, Tom

APPLICANT: BATES, Stephen

APPLICANT: FANTIN, Dennis

TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS

FILE REFERENCE: MM1100-2

CURRENT APPLICATION NUMBER: US/10/750,185

CURRENT FILING DATE: 2003-12-31

PRIOR APPLICATION NUMBER: US 60/437,482

PRIOR FILING DATE: 2002-12-31

NUMBER OF SEQ ID NOS: 64922

SOFTWARE: Patent In version 3.1

SEQ ID NO 57746

LENGTH: 747

TYPE: DNA

ORGANISM: Bovine

US-10-750-185-57746

Query Match 2.0%; Score 34.2; DB 6; Length 747;

Best Local Similarity 55.5%; Pred. No. 3.1;

Matches 66; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

Qy 1275 TGTGCTATTTCAGTACAGAGATCACTAGACTTTTGAATAATTCGACCGCAATTTGGCAT 1334

Db 368 TTTATGTTTGAACACAGACACTAAATGACATTTTATATGTCGATCGACAGTTTGTCT 309

Qy 1335 TTTTGAAGAATGGAAGGATGTTACCGCTGACATATTAAGGAATAGTAGTTTTC 1393

Db 308 TTTATTTGTGAATATGATTAATCTTGTCTTTTCAAGGGTTTAGGTAGACTGGCTTTC 250

<p>RESULT 7</p> <p>US-11-117-187-201</p> <p>; Sequence 201, Application US/11117187</p> <p>; Publication No. US20050266560A1</p> <p>; GENERAL INFORMATION:</p> <p>; APPLICANT: PREUSS, DAPHNE</p> <p>; APPLICANT: COPENHAVER, GREGORY</p> <p>; TITLE OF INVENTION: PLANT ARTIFICIAL CHROMOSOME COMPOSITIONS AND METHODS</p> <p>; FILE REFERENCE: ARCD:309US</p> <p>; CURRENT APPLICATION NUMBER: US/11/117,187</p> <p>; PRIOR FILING DATE: 2005-04-28</p> <p>; PRIOR APPLICATION NUMBER: US/09/531,120</p> <p>; PRIOR FILING DATE: 2000-03-17</p> <p>; PRIOR APPLICATION NUMBER: 60/125,219</p> <p>; PRIOR FILING DATE: 1999-03-18</p> <p>; NUMBER OF SEQ ID NOS: 212</p> <p>; SOFTWARE: PatentIn Ver. 2.1</p> <p>; SEQ ID NO 201</p> <p>; LENGTH: 80450</p> <p>; TYPE: DNA</p> <p>; ORGANISM: Arabidopsis thaliana</p> <p>US-11-117-187-201</p>		<p>Query Match</p> <p>Best Local Similarity 2.0%; Score 34.2; DB 7; Length 80450;</p> <p>Mismatches 69; Conservative 0; Mismatches 58; Indels 0; Gaps 0;</p>		<p>QY 1087 TTGGGCGAGTTTTCACGAGTATCTTGAGCCATTAACTAAATGATGTCCAGGTTGAT 1146</p> <p>DB 13727 TAGGTTAATTTTACCAGCAATTAATCTTAGAGACAAATGAAGGTTCAAGTGGAG 13786</p> <p>QY 1147 TGAAGTCAATGGACCTTAGTACCTTTTGAGGACCAATACGTGAAACACTTTGGTGAC 1206</p> <p>DB 13787 TCAAAGCAAGGAGCGGAGTCACCTTATTATGGGAATATGAGTGTGGCACTTTTTCCTCC 13846</p> <p>QY 1207 TTGGTTA 1213</p> <p>DB 13847 TTATTTA 13853</p>		<p>RESULT 8</p> <p>US-10-750-185-33227</p> <p>; Sequence 33227, Application US/10750185</p> <p>; Publication No. US20050260603A1</p> <p>; GENERAL INFORMATION:</p> <p>; APPLICANT: MMI GENOMICS, INC.</p> <p>; APPLICANT: DENISE, Sue K.</p> <p>; APPLICANT: KERR, Richard</p> <p>; APPLICANT: ROSENFELD, David</p> <p>; APPLICANT: HOLM, Tom</p> <p>; APPLICANT: BATES, Stephen</p> <p>; APPLICANT: FANTIN, Dennis</p> <p>; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS</p> <p>; FILE REFERENCE: MM1100-2</p> <p>; CURRENT APPLICATION NUMBER: US/10/750,185</p> <p>; CURRENT FILING DATE: 2003-12-31</p> <p>; PRIOR APPLICATION NUMBER: US 60/437,482</p> <p>; PRIOR FILING DATE: 2002-12-31</p> <p>; NUMBER OF SEQ ID NOS: 64922</p> <p>; SOFTWARE: PatentIn version 3.1</p> <p>; SEQ ID NO 33227</p> <p>; LENGTH: 1276</p> <p>; TYPE: DNA</p> <p>; ORGANISM: Bovine</p> <p>US-10-750-185-33227</p>		<p>Query Match</p> <p>Best Local Similarity 2.0%; Score 34; DB 6; Length 1276;</p> <p>Mismatches 76; Conservative 0; Mismatches 70; Indels 0; Gaps 0;</p>		<p>QY 942 AAATCTACTGGGACGAATTTATCTCAAAGTGCCAAAGGCTTACTGGGACGACTGGCT 1001</p>	
<p>Db 1088 ATAAGCTCATCTAGACTTAACTTCTCCAAAGTGCCACAGTAATTCAGATTTCTTTGCT 1147</p> <p>QY 1002 AGACTCAAAGAGAAATCACAGAGTTCGACAAATTTATTCGCCCAAGAGTTTTCGACACATA 1061</p> <p>Db 1148 AAGATTAAACACACTTCTTTCAGACAAACACATATTTCTGTACCTATAAATTTTGTGCATACA 1207</p> <p>QY 1062 TAATTTTGGTGAGCATGGTTCTAGTT 1087</p> <p>Db 1208 TTGCGTGAGTGTCCTCCCGATGCAAAAT 1233</p>		<p>RESULT 9</p> <p>US-10-793-626-4157/c</p> <p>; Sequence 4157, Application US/10793626</p> <p>; Publication No. US20050255478A1</p> <p>; GENERAL INFORMATION:</p> <p>; APPLICANT: KIMMERLY, WILLIAM JOHN</p> <p>; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS</p> <p>; FILE REFERENCE: PU3480US</p> <p>; CURRENT APPLICATION NUMBER: US/10/793,626</p> <p>; CURRENT FILING DATE: 2004-03-04</p> <p>; PRIOR APPLICATION NUMBER: 60/164,258</p> <p>; PRIOR FILING DATE: 1999-11-09</p> <p>; NUMBER OF SEQ ID NOS: 4472</p> <p>; SOFTWARE: PatentIn Ver. 2.1</p> <p>; SEQ ID NO 4157</p> <p>; LENGTH: 3619</p> <p>; TYPE: DNA</p> <p>; ORGANISM: Artificial Sequence</p> <p>; FEATURE:</p> <p>; OTHER INFORMATION: Description of Artificial Sequence: synthetic</p> <p>; OTHER INFORMATION: nucleic acid sequence</p> <p>US-10-793-626-4157</p>		<p>Query Match</p> <p>Best Local Similarity 2.0%; Score 34; DB 6; Length 3619;</p> <p>Mismatches 85; Conservative 0; Mismatches 85; Indels 0; Gaps 0;</p>		<p>QY 1213 AAAAGGCTAAGCCCATCCATGGAGCTGATGCTGCTTGAAGCAATTAACATAGATGGT 1272</p> <p>Db 2960 ACAATCCTGCACCTAGACTCAGAAAGGTAATCTGTATGGTAACTATTATGGAACAAGAC 2901</p> <p>QY 1273 GATGTCGGTATTCAGTACAGAGATCAACTAGACTTTGAAATATCCACGGCAATTTGGC 1332</p> <p>Db 2900 GATATGACGATTGATAGTGGATTTTATCAAAACACCTAAATACAGCTTAGGGAACATGTA 2841</p> <p>QY 1333 ATTTTGAAGAATGGAAGGATGGTGATCCACCGTGACGATATAAAGGAAT 1382</p> <p>Db 2840 TGGTATGACACTAATAAAGATGGTATTCAGGTGATGATCAAAAAGGAAT 2791</p>		<p>RESULT 10</p> <p>US-10-995-561-13293</p> <p>; Sequence 13293, Application US/109955561</p> <p>; Publication No. US20050272054A1</p> <p>; GENERAL INFORMATION:</p> <p>; APPLICANT: CARGILL, Michele et al.</p> <p>; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH</p> <p>; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF</p> <p>; TITLE OF INVENTION: DETECTION AND USES THEREOF</p> <p>; FILE REFERENCE: CL001559</p> <p>; CURRENT APPLICATION NUMBER: US/10/995,561</p> <p>; CURRENT FILING DATE: 2004-11-24</p> <p>; NUMBER OF SEQ ID NOS: 85702</p> <p>; SOFTWARE: FastSeq for Windows Version 4.0</p> <p>; SEQ ID NO 13293</p> <p>; LENGTH: 645179</p> <p>; TYPE: DNA</p> <p>; ORGANISM: Homo sapiens</p> <p>US-10-995-561-13293</p>		<p>Query Match</p> <p>2.0%; Score 34; DB 6; Length 645179;</p>	

Best Local Similarity 46.5%; Pred. No. 2e+02;  
Matches 106; Conservative 1; Mismatches 121; Indels 0; Gaps 0;  
QY 1165 AGTACCTTTGGAGGCAATTACGTGAAACACCTTTGGTGACTGGTTAAAGAGCTTAAG 1224  
DB 149618 ATTTACATGGGAAGTCTACTGTGTACCAAGTTGTCATATTTTGTGTAAAGTTCTCAG 149677  
QY 1225 CCCATCCATGAGCTGATGCTCTTGAAGCATTTTACATAGATGGTGTGCGTATT 1284  
DB 149678 CCATGGTATGATGCAATTTGGTGTGTGACGCCATCTCTCAGACCTTTATTTCTCTGTT 149737  
QY 1285 CAGTACAGAGATCAACTAGACTTTTGAAATATCGCAGCGCAATTTGGCAATTTTGAAGAA 1344  
DB 149738 CAGCCATTAACAAATGCAATTTTCAATGTTCAGGTAGTTTCAATTTTAAAGATA 149797  
QY 1345 TGAAGAGATGGTATACCAAGTGCAGCATATTAAGGAATAGTAGTTTC 1392  
DB 149798 CAGAAACTAGTATTCAATGGTTAAATAAAGCAATACCTATC 149845

RESULT 11  
US-10-750-185-52662  
; Sequence 52662, Application US/10750185  
; Publication No. US20050260603A1  
; GENERAL INFORMATION:  
; APPLICANT: MMI GENOMICS, INC.  
; APPLICANT: DENISE, Sue K.  
; APPLICANT: KERR, Richard  
; APPLICANT: ROSENFELD, David  
; APPLICANT: HOLM, Tom  
; APPLICANT: BATES, Stephen  
; APPLICANT: FANTIN, Dennis  
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS  
; FILE REFERENCE: MM1100-2  
; CURRENT APPLICATION NUMBER: US/10/750,185  
; CURRENT FILING DATE: 2003-12-31  
; PRIOR APPLICATION NUMBER: US 60/437,482  
; PRIOR FILING DATE: 2002-12-31  
; NUMBER OF SEQ ID NOS: 64922  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 52662  
; LENGTH: 1056  
; TYPE: DNA  
; ORGANISM: Bovine 19866881397479  
US-10-750-185-52662

Query Match 1.9%; Score 33.8; DB 6; Length 1056;  
Best Local Similarity 56.9%; Pred. No. 5;  
Matches 62; Conservative 0; Mismatches 47; Indels 0; Gaps 0;  
QY 142 TTTTGTGTGATTTCCGGTACCTCTCCTCATCTTGGCTGTGCTGCTTACATCTACACAG 201  
DB 370 TTTGGCAGACATTTCCCT 429  
QY 202 ATGCGCTTTTTCGACAGTGCAGATATGAGATCGCTTGTGCTG 250  
DB 430 TTAGAAGGATTTGCTTTCGAGAAGAAATGAACCGCTCTTTGCTTCTG 478

RESULT 12  
US-10-793-626-4439/C  
; Sequence 4439, Application US/10793626  
; Publication No. US20050255478A1  
; GENERAL INFORMATION:  
; APPLICANT: KIMMERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PU3480US  
; CURRENT APPLICATION NUMBER: US/10/793,626  
; CURRENT FILING DATE: 2004-03-04  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 4439  
; LENGTH: 2460  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: nucleic acid sequence  
US-10-793-626-4439

Query Match 1.9%; Score 33.8; DB 6; Length 2460;  
Best Local Similarity 44.9%; Pred. No. 8.5;  
Matches 170; Conservative 0; Mismatches 207; Indels 2; Gaps 1;  
QY 884 AAGATCCTTATGCTCTTTACCGCTCAGATTTTTCOCGGTCTTGGATGGATGCTTTCAA 943  
DB 1189 AAAATTTATPACCAATAACATAGATACCTTTTTCATTAAGTTTCACTAGTTTTCAAA 1130  
QY 944 AATCTACTTGGGACGAATTTCTCAAGTGGCCAAAGGCTTACTGGGACGACTGGCTAA 1003  
DB 1129 AAAATTTTATAATTTTATGATACTTAAAGTGATATTTTCTTACATATATATTTAA 1070  
QY 1004 GACTCAAGAGAAATCACAGAGGTGACAAATTTTTCGCCCAAGAGTTTGCAGAACATATA 1063  
DB 1069 TTTTCAAGGAGATGTAAAAAGTTGAAAAATTTTCGCAAACTAATTTTGGTGGCATTTT 1010  
QY 1064 ATTTTGGTGTAGCATGGTTCTTAGTTTGGGGCAGTTTTCGAAGCAGTATCTTGAGCCAA 1123  
DB 1009 AGTATCGGGTTCAGGATAGCGAGTGTACAAACAAATATACTCAGCCAAAGAAAGTCA 950  
QY 1124 --AATAAATGATGTCAGGTTGATTTGGAAGTCAATGAGACCTTAGTTACCTTTTGGAGA 1181  
DB 949 CGATTCAACTCTCAAAATATTAATTTAGTGGGAACGTATGATATCTTCTCAAGTTGATTC 890  
QY 1182 CAATTACGTGAACACATTTTGGTGTGCTTGTAAAGAGCTAAGCCCATCCATGGAGCTGA 1241  
DB 889 CAAACGATGAACAAATTTTAAAGAAATAGAAAAAGAGATATAATTTTCCACATAACTAA 830  
QY 1242 TGCTGTCTTGAAAGCATTT 1260  
DB 829 ACATGGAATTAAGTCGTT 811

RESULT 13  
US-10-793-626-3709  
; Sequence 3709, Application US/10793626  
; Publication No. US20050255478A1  
; GENERAL INFORMATION:  
; APPLICANT: KIMMERLY, WILLIAM JOHN  
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS  
; FILE REFERENCE: PU3480US  
; CURRENT APPLICATION NUMBER: US/10/793,626  
; CURRENT FILING DATE: 2004-03-04  
; PRIOR APPLICATION NUMBER: 60/164,258  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 4472  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3709  
; LENGTH: 2683  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: synthetic  
; OTHER INFORMATION: nucleic acid sequence  
US-10-793-626-3709

Query Match 1.9%; Score 33.8; DB 6; Length 2683;  
Best Local Similarity 44.9%; Pred. No. 8.9;  
Matches 170; Conservative 0; Mismatches 207; Indels 2; Gaps 1;  
QY 884 AAGATCCTTATGCTCTTTACCGCTCAGATTTTTCOCGGTCTTGGATGGATGCTTTCAA 943  
DB 37 AAAATTTATTAACCAATAACATAGATACCTTTTTCATTAAGTTTCACTAGTTTTCAAA 96

```
Qy 944 AATCTACTGGGACGAATTTCTCCAAAGTGGCCAAAGGCTTACTGGGACGACTGGCTAA 1003
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 97 AAAATTTTATAATTTATGATACTTAAAGTGTATATTTTCTACATAATATTTAA 156
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1004 GACTCAAAGAGATCACAGAGGTGCACAAATTTATTCGCCCCAGAAAGTTTGCAGAACATATA 1063
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 157 TTTTCAAGGAGATGATAAAAGTTGAAAATTTTCGCAAACTAATTTTGGTGGCAITTT 216
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1064 ATTTTGGTGAGCATGGTTCTAGTTTGGGGCAGTTTTTCAAGCAGATATCTTGAGCCAATTA 1123
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 217 AGTATCGGGTCAGGATAGCGAGTGTACAAACAAATATAACTCACGCAAAAGAAAGTCA 276
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1124 --AACTAAATGATCTCAGAGTTGATTCGAACTCAATGACCTTAGTTACCTTTTGGAGGA 1181
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 277 CGATTCACTCTCAAAATATTAATTAGTGGGAACGTATGATACTTCTCAAGTTGATTC 336
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1182 CAATTAGCTGAACACATTTTGGTGACTTTGGTTAAAGAGCTTAAGCCCATCCATGAGCTGA 1241
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 337 CAAAACGATGAACAAATTTAAAGAAATAGAAAAGAAAGATAATTAATTTCCACATACTAA 396
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1242 TGCTGTCTTGAAGCATTT 1260
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 397 ACATGGAATAAAGTCGTT 415
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

## RESULT 14

```
US-10-793-626-3969/c
; Sequence 3969, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10793.626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3969
; LENGTH: 3022
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-3969
```

```
Query Match 1.9%; Score 33.8; DB 6; Length 3022;
Best Local Similarity 44.9%; Pred. No. 9.6;
Matches 170; Conservative 0; Mismatches 207; Indels 2; Gaps 1;

Qy 884 AAGATCCTTATGCTCTTTACCGCTCAGATTTTTTCCCGGCTCTGGGATGGATGCTTTCAA 943
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1152 AAAATTTATACCAATAACATAGATACTTTTTTCATTAAGTTTCACTAGITTTTCAA 1093
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 944 AATCTACTGGGACGAATTTATCTCCAAAGTGGCCAAAGGCTTACTGGGACGACTGGCTAA 1003
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1092 AAAATTTTATAATTTATGATACTTAAAGTGTATATTTTTTCTACATAATATTTAA 1033
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1004 GACTCAAAGAGATCACAGAGGTGCACAAATTTATTCGCCCCAGAAAGTTTGCAGAACATATA 1063
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1032 TTTTCAAGGAGATGATAAAAGTTGAAAATTTTCGCAAACTAATTTTGGTGGCAITTT 973
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1064 ATTTTGGTGAGCATGGTTCTAGTTTGGGGCAGTTTTTCAAGCAGATATCTTGAGCCAATTA 1123
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 972 AGTATCGGGTCAGGATAGCGAGTGTACAAACAAATATAACTCACGCAAAAGAAAGTCA 913
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1124 --AACTAAATGATCTCAGGTTGATTCGAACTCAATGACCTTAGTTACCTTTTGGAGGA 1181
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 912 CGATTCACTCTCAAAATATTAATTAGTGGGAACGTATGATACTTCTCAAGTTGATTC 853
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1182 CAATTACGTGAACACATTTTGGTGACTTTGGTTAAAGAGCTTAAGCCCATCCATGAGCTGA 1241
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Db 852 CAAAACGATGAACAATTTTAAGAAATAGAAAAGAGATAATAATTTCCACATACTAA 793
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1242 TGCTGTCTTGAAGCATTT 1260
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 792 ACATGGAATAAAGTCGTT 774
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

## RESULT 15

```
US-10-793-626-3361
; Sequence 3361, Application US/10793626
; Publication No. US20050255478A1
; GENERAL INFORMATION:
; APPLICANT: KIMMERLY, WILLIAM JOHN
; TITLE OF INVENTION: STAPHYLOCOCCUS EPIDERMIDIS NUCLEIC ACIDS AND PROTEINS
; FILE REFERENCE: PU3480US
; CURRENT APPLICATION NUMBER: US/10793.626
; CURRENT FILING DATE: 2004-03-04
; PRIOR APPLICATION NUMBER: 60/164,258
; PRIOR FILING DATE: 1999-11-09
; NUMBER OF SEQ ID NOS: 4472
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3361
; LENGTH: 4039
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic
; OTHER INFORMATION: nucleic acid sequence
US-10-793-626-3361
```

```
Query Match 1.9%; Score 33.8; DB 6; Length 4039;
Best Local Similarity 44.9%; Pred. No. 12;
Matches 170; Conservative 0; Mismatches 207; Indels 2; Gaps 1;

Qy 884 AAGATCCTTATGCTCTTTACCGCTCAGATTTTTTCCCGGCTCTGGGATGGATGCTTTCAA 943
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1434 AAAATTTATATACCAATAACATAGATACTTTTTTCATTAAGTTTCACTAGITTTTCAA 1493
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 944 AATCTACTGGGACGAATTTATCTCCAAAGTGGCCAAAGGCTTACTGGGACGACTGGCTAA 1003
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1494 AAAATTTTATAATTTATGATACTTAAAGTGTATATTTTTTCTACATAATATTTAA 1553
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1004 GACTCAAAGAGATCACAGAGGTGCACAAATTTATTCGCCCCAGAAAGTTTGCAGAACATATA 1063
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1554 TTTTCAAGGAGATGTAAAAAAGTTGAAAAATTTTCGCAAACTAATTTTGGTGGCAITTT 1613
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1064 ATTTTGGTGAGCATGGTTCTAGTTTGGGCGAGTTTTTCAAGCAGATATCTTGAGCCAATTA 1123
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1614 AGTATCGGGTCAGGATAGCGAGTGTACAAACAAATATTAATCTACGCAAAAGAAAGTCA 1673
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1124 --AACTAAATGATGTCCAGGTTGAITTTGGAAGTCAATTTGGAACCTTTAGTTACCTTTTGGAGGA 1181
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1674 CGATTCACTCTCAAAATATTAATAATTTAGTGGGAACGTATGATACTTCTCAAGTTGATTC 1733
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1182 CAATTACGTGAACACATTTTGGTGACTTTGGTTAAAGAGCTTAAGCCCATCCATGAGCTGA 1241
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1734 CAAAACGATGAACAATTTAAAGAAATAGAAAAGAGATAATAATTTCCACATACTAA 1793
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1242 TGCTGTCTTGAAGCATTT 1260
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1794 ACATGGAATAAAGTCGTT 1812
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Search completed: December 14, 2005, 13:35:44  
Job time : 295 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2005, 08:02:40 ; Search time 6397 Seconds  
(without alignment)  
12704.262 Million cell updates/sec

Title: US-10-615-144-3  
Perfect score: 1737  
Sequence: 1 gaattcgccgcattga.....ataatcgccgcgaattc 1737

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues  
Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:  
1: gb\_est1:  
2: gb\_est2:  
3: gb\_est3:  
4: gb\_hc:  
5: gb\_est4:  
6: gb\_est5:  
7: gb\_est6:  
8: gb\_est7:  
9: gb\_gss1:  
10: gb\_gss2:  
11: gb\_gss3:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	739.6	42.6	894	DN981229	DN981229 SV6_35B12
2	723.4	41.6	925	BQ505949	BQ505949 EST613364
3	722.4	41.6	1654	CNS0A4J5	BX828572 Arabidops
4	630	36.3	768	EG594788	BG594788 EST493466
5	599.6	34.5	772	BQ505948	BQ505948 EST613363
6	593	34.1	724	B1179808	B1179808 EST520753
7	580.2	33.4	642	BG592816	BG592816 EST491494
8	578.6	33.3	643	AI486771	AI486771 EST515723
9	553.6	31.9	661	BG889872	BG889872 EST515723
10	537.8	31.0	673	AW979500	AW979500 EST310521
11	527	30.3	659	AW735827	AW735827 EST336595
12	517.2	29.8	843	BM779745	BM779745 EST590321
13	514	29.6	574	AW931180	AW931180 EST357023
14	510.6	29.4	758	AJ806862	AJ806862 AJ806862
15	508.2	29.3	852	BM780056	BM780056 EST590632
16	496.6	28.6	874	DR927057	DR927057 EST111859
17	484.8	27.9	1036	BX839862	BX839862 EX839862
18	465.6	26.8	875	DR015231	DR015231 STRS1_2_F
19	456.8	26.3	705	CO491834	CO491834 G.h.fbr-s
20	453.8	26.1	837	DR692299	DR692299 EST108238
21	445.6	25.7	674	BU837218	BU837218 T096B09 P
22	444.4	25.6	454	BP527408	BP527408 BP527408

23	438.8	25.3	681	1	AI725875
24	438.4	25.2	640	6	CA927022
25	431.4	24.8	849	6	CB631083
26	428.6	24.7	703	5	BQ858509
27	411.2	23.7	800	6	CB67281
28	407	23.4	619	3	BP956906
29	404.6	23.3	844	8	CV763947
30	398.8	23.0	555	7	CN494603
31	398	22.9	664	2	BE805530
32	393.4	22.1	484	3	BP901768
33	382.8	22.0	629	2	BE426174
34	379.8	21.9	716	7	CK861930
35	374.4	21.6	831	8	DR927051
36	372.2	21.4	591	2	EG409413
37	361	20.8	755	7	CV534697
38	358.2	20.6	574	7	CV457771
39	348	20.0	477	8	DR909675
40	343.8	19.8	530	1	AW618699
41	341.2	19.6	627	7	CV503179
42	338.4	19.5	615	8	CK172334
43	338.2	19.5	519	7	CO088227
44	335.2	19.3	633	1	AJ805159
45	333.6	19.2	736	6	CA268807

ALIGNMENTS

RESULT 1  
DN981229  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

DN981229 894 bp mRNA linear EST 15-JUN-2005  
SV6\_35B12\_SV6\_35T\_093\_SV6 Solanum chacoense cDNA, mRNA sequence.  
DN981229  
DN981229.1 GI:67769040  
EST.  
Solanum chacoense (Chaco potato)  
Solanum chacoense  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
asterids; lamiales; Solanales; Solanaceae; Solanum.  
1 (bases 1 to 894)  
Germain,H., Rudd,S., Zotti,C., Caron,S., O'Brien,M., Chantha,S.-C.,  
Lagace,M., Major,F. and Matton,D.P.  
A 6374 unigene set corresponding to low abundance transcripts  
expressed following fertilization in Solanum chacoense Bitt., and  
expression profile of 30 receptor-like kinases  
Unpublished (2005)  
Contact: Daniel P. Matton  
Daniel P. Matton  
Universite de Montreal  
Institut de recherche en biologie vegetale, 4101 rue Sherbrooke  
est, Montreal, QC, Canada, H1X 2B2  
Tel: 5148723967  
Fax: 5148729406  
Email: dp.matton@umontreal.ca.

FEATURES  
source

Location/Qualifiers  
1..894  
/organism="Solanum chacoense"  
/mol\_type="mRNA"  
/db\_xref="taxon:4108"  
/tissue\_type="2-6 days post-pollination ovaries"  
/dev\_stage="Zygotic to globular embryo"  
/lab\_host="SOLR"  
/clone\_lib="SV6"  
/note="Organ: Deparicarped ovary/ovules with placenta  
tissue; Vector: pBS SK; Site 1: EcoRI; Site 2: XhoI;  
Deparicarped ovaries were isolated from a compatible cross  
between self-incompatible accessions having the S12S14  
self-incompatibility alleles (G4 line as female) and  
S11S13 self-incompatibility alleles (V22 line as male).  
Those two lines were derived from the parental lines PI  
458314 (which carries the S11 and S12 self-incompatibility  
alleles) and line PI 230582 (which carries the S13 and

S14 alleles) obtained through the Potato Introduction Station (Sturgeon Bay, WI). Plants were grown in greenhouses under natural light, fruit collection took place in the morning during summer (approx. 14h daylight), and fruits were pooled from different plants. Total RNA was isolated as described in EMBO J. 4: 2411-2418, 1985. The poly A+ RNA was purified from the total RNA by affinity chromatography using oligo dt agarose beads type VII (Amersham). Mean clone length of the library was 1,5 kb."

ORIGIN		Query Match	42.6%;	Score	739.6;	DB	8;	Length	894;
		Best Local Similarity	90.8%;	Pred. No.	8.2e-200;				
		Matches	813;	Conservative	0;	Mismatches	77;	Indels	5;
						Gaps	3;		
Qy	457	GTGCTGTTATGGCTTGCATCGCGCTGATTACCTGGAAAGACTATTAATCATCTTAAAA	516						
Db	5	GTGCTGTTATGGCTTGCAGTCGCTGACTACCTGGAGAGACTATTAATCATCTTAAAA	64						
Qy	517	TACCAATATCTGTTGGCTCAAAATATCTCTTTTCATATCCAGGATGGATCATCTCT	576						
Db	65	TACCAAAATCTGTTGCATCAAAATATCTCTTTTCATATCCAGGATGGATCAATCTCT	124						
Qy	577	GATGTCAGGAGCTTGCTTTGAGCTATCATCAGCTGAGCTATATGACGACCTTGGATTTT	636						
Db	125	GATGTTAGAAAGCTTGCTTTGAGCTATCATCAGCTGACATATATGACGACCTTGGATTTT	184						
Qy	637	GAACTGTGCATCTGAAAGACGAGGAGCTGATTGCATCTACTACAAATTTGCACGTCAT	696						
Db	185	GAACTGTGCATCTGAAAGACGAGGAGCTGTTGCATCTACTACAAATTTGCACGTCAT	244						
Qy	697	TACAAGTGGGCAATGGATCAGCTGTTTACAAGCATAAATTTTACCGCTGTTTACATCTA	756						
Db	245	TACAAGTGGGCAATGGATCAGCTGTTTACAAGCATAAATTTTACCGCTGTTTACATCTA	304						
Qy	757	GAGATGATATGGAATGGCCCTGATTTTTCACATTTTTCAGCTTTTGGGCTGGAGCTCTT	816						
Db	305	GAGATGATATGGAATTTGCTGCTGATTTTTCAGCTTTTTCAGCTTTTGGGCTGGAGCTCTT	364						
Qy	817	CTTGACAGAGACAAAGTCGATTATGGCTATTTCTTTTGGAAATGACAAATGGACAAATGCAG	876						
Db	365	CTTGACAGAGACAAAGTCGATTATGGCTATTTCTTTTGGAAATGACAAATGGACAAAGCGAG	424						
Qy	877	TTTGTCAGAGATCCTTATGCTCTTTACCGCTCAGATTTTTTTCCTGGCTTTGGATGGATG	936						
Db	425	TTTGTCAGAGATCCTTATGCTCTTTACCGCTCAGATTTTTTTCCTGGCTTTGGATGGATG	484						
Qy	937	CTTTCAAAATCTACTTGGGACGAATATCTCCAAAGTGGCCAAAGGCTTACTGGGACGAC	996						
Db	485	CTTTCAAAATCAACTTGTTCGAACTATCTCCAAAGTGGCCAAAGGCTTACTGGGATGAC	544						
Qy	997	TGGCTAAGACTCAAGAGAAATCACAGAGGTCGACAAATTTATTCGCCCCAGAAAGTTTGCGA	1056						
Db	545	TGGCTAAGGCTCAAGAGAAATCACAGAGGTCGACAAATTTATTCGCCCCAGAAAGTTTGCGA	604						
Qy	1057	ACATATAATTTTGTGAGCATGGTTCTAGTTTGGGCGAGTTTTTTTCAAGCATATCTTTGAG	1116						
Db	605	ACATATAATTTTGTGAGCATGGCTCTAGTTTGGGCGAGTTTTTTTAAAGCATATCTTCGAG	664						
Qy	1117	CCAATTAACCTAAATGATCCAGGTTGATTTGGAGTCAATGGACCTTAGTTACCTTTTG	1176						
Db	665	CCAATTAACCTAAATGATGTCAGGTTGATTTGGAGTCAATGGACCTTAGTTACCTTTTG	724						
Qy	1177	GAGGACAAATACGTGAAACACTTTTGGTGACCTTGTGTTTAAAGAGGCTAAGCCCATCCATGGA	1236						
Db	725	GAGGACAAATACGTGAAACACTTTTGGTGACCTTGTGTTTAAAGAGGCTAAGCCCATCCATGGA	783						
Qy	1237	GCTGATGCTGTTGAAAGCAATTTAACATAGATGGTGATGTCGCTATTTACGATCAGAGAT	1296						
Db	784	GCTGATGCTGNTTTGAAAGC-TTTTAAATATATGNGATGGGCGCTNTTCGG--NCANAN	839						
Qy	1297	CAACTAGACTTTTGAAATATCGCAGCGCAATTTTGGCATTTTGTGAAGATGGAAGG	1351						

Db	840	ACACCTANACTTGGAAATTCNCTCGCAGCTGGCATTTTGAANNAGAGGGG	894
RESULT 2			
BO505949/c			
LOCUS	BO505949	925 bp	mRNA linear EST 07-MAR-2003
DEFINITION	EST613364 Generation of a set of potato cDNA clones for microarray analyses mixed potato tissues Solanum tuberosum cDNA clone STMGH33 3' end, mRNA sequence.		
ACCESSION	BO505949		
VERSION	BO505949.2	GI:21921856	
KEYWORDS	EST.		
SOURCE	Solanum tuberosum (potato)		
ORGANISM	Solanum tuberosum		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamiids; Solanales; Solanaceae; Solanum.		
	1 (bases 1 to 925)		
AUTHORS	Buell, C.R., Hart, A., Baker, B., Tanksley, S., Fry, W., Smart, C., Restrepo, S., Griffiths, H., van der Hoeven, R., Tsai, J. and Karanycheva, S.A.		
	Generation of a set of potato cDNA clones for microarray analyses		
TITLE	Unpublished (2002)		
JOURNAL	On Jun 10, 2002 this sequence version replaced gi:21364818.		
COMMENT	Other ESTs: EST613363 Contact: Robin Buell The Institute for Genomic Research 9712 Medical Center Dr, Rockville, MD 20850, USA Email: potato-array@tigr.org This clone can be obtained from the University of Arizona Genomics Institute. Orders can be made through URL: http://genome.arizona.edu/orders/ Seq primer: T7.		
FEATURES	Location/Qualifiers		
source	1..925		
	/organism="Solanum tuberosum"		
	/mol_type="mRNA"		
	/cultivar="Kennebec or Binjite"		
	/db_xref="taxon:4113"		
	/clone="STMGH33"		
	/tissue_type="mixed tissues"		
	/lab_host="SOLR"		
	/clone_lib="Generation of a set of potato cDNA clones for microarray analyses mixed potato tissues"		
	/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2: XhoI; supplier: Combination of untreated and Phytophthora infestans-treated libraries of stolons, leaves, leaflets, axillary buds of stem explants, petioles, germinating eyes, tubers, or roots."		
ORIGIN			
Query Match	41.6%;	Score 723.4;	DB 5; Length 925;
Best Local Similarity	88.4%;	Pred. No. 3.6e-195;	
Matches 811; Conservative	0;	Mismatches 96;	Indels 10; Gaps 2;
QY	703	TGGGCATTTGGATCAGCTGTTTACAGCATAAATTTAGCCGTGTTATCATAGAGAT	762
Db	925	TGGGCATTTGGATCAGCTGTTTACAGCATAAATTTAGCCGTGTTATCATAGAGAT	866
QY	763	GATATGGAATTCGCCCTGATTTTGTGCTTTTGTAGCTTTTGTAGGCTGAGCTACTCTCTTGAC	822
Db	865	GATATGGAATTCGCTGCTGATTTTGTGCTTTTGTAGCTTTTGTAGGCTGAGCTACTCTCTTGAC	806
QY	823	AGAGACAAAGTCGATTATGGCTATTTCTTTTGAATGCAATGCAAAATGCAGTTTGTGTC	882
Db	805	AGAGACAAAGTCGATTATGGCTATTTCTTTTGAATGCAATGCAAAAGCGAGTTCGTC	746
QY	883	CAAGATCCTTATGCTCTTTACCGCTCAGATTTTTTCCCGGTCTTGGATGGATGCTTTCA	942
Db	745	CAAGATCCTGATGCTCTTTACCGCTCAGACTTTTTTCTGGTCTTGGATGGATGCTTTCA	686
QY	943	AAATCTACTTGGGACGAATTTATCTCCAAAGTGGCCAAAGGCTTACTGGGACGACTGGCTA	1002



```
Db 685 AATCACTTGGTCGAACATATCCAAAGTGGCCAAAGGCTTACTGGATGACTGGCTA 626
Qy 1003 AGACTCAAGAGATCAAGAGGTCGACAAATTTATTCGCCAGAGAGTTTGGAGAACATAT 1062
Db 625 AGGCTGAAAGAAAATCAAGAGGTCGACAAATTTATTCACCCAGAGAGTTTGGAGAACGTCAC 566
Qy 1063 AATTTTGGTGAAGCATGGTCTAGTTTGGGCGAGTTTTCAGAGCATATCTTGGAGCAATT 1122
Db 565 AATTTTGGTGAAGCATGGTCTAGTTTGGGCGAATTTTTCAGAGCATATCTTGGAGCAATT 506
Qy 1123 AAACATAATGATGTCAGAGTTGATTTGAAGTCAATGACCTTTCAGTTTGGAGGAC 1182
Db 505 AAGCTAATGATGTCAGAGTTGATTTGAAGTCAATGACCTTTCAGTTTGGAGGAC 446
Qy 1183 AATACGTGAACACTTTGGTGACCTTGGTTAAAGAGGCTAAGCCCATCCATGGAGCTGAT 1242
Db 445 AACTATGTGAACACTTTGGTGACCTTGGTTAAAGAGGCTAAGCCCATCCATGGAGCTGAT 386
Qy 1243 GCTGTCTTGAAGACATTTAAACATAGATGCTGATGCGTATTTCAGTACAGAGATCAACTA 1302
Db 385 GCTGTCTTGAAGACATTTAAACATAGATGCTGATGCGTATTTCAGTACAGAGATCAACTA 326
Qy 1303 GACTTTGAAAATATCGCAGGCAATTTGGCAATTTTGAAGATGGAAGATGGTGTACCA 1362
Db 325 GACTTTGAAAGCTATCGCTTGACAGTTTGGCAATTTTGAAGATGGAAGATGGTGTACCA 266
Qy 1363 COTGAGCATATAAGGAATAGTAGTTTCCGGTACCAAGCTCCAGAGCTGATTCCTT 1422
Db 265 CGGCGAGCATATAAAGGATAGTAGTTTCCGGTTCACCAATCTAGAGCTGTTCTCTT 206
Qy 1423 GTTGCCATGATTCGCTTCAACACTCGGAATTTGAAGATGATCTTAAAGATGATGATGC 1482
Db 205 GTTGCCCTGATCTCTTGAACACTTGGAGTTGAAGATGATCTTGAAGATGATGATGAT 148
Qy 1483 AGGAGCCCGGCAAAATTTTGACTTATTGGGTAGGATGATCGAGCTGACATAAACA 1542
Db 147 -GGAGCTGAGCAACAAATTTAGACTTATTGGTAGGATACATTTGAAAGAGCTGACCGA 89
Qy 1543 TGA-----TTTACCAGTACATACAAAGCTTTTAAAGTATCGAGGAGCTCACTGT 1595
Db 88 AAGATGATGACTTCCAGTAGGTACATGCAACATTTTAAATGTTAATGGAAGGAAACCCCTGC 29
Qy 1596 TCTAGTGTGAAGGAT 1612
Db 28 TTATGTTGGAATGGAT 12

RESULT 3
CNS0A4J5 1654 bp mRNA linear HTC 06-FEB-2004
LOCUS Arabidopsis thaliana Full-length cDNA complete sequence from clone
DEFINITION GSL7S1182D01 of Silique of strain col-0 of Arabidopsis thaliana
(thale cress).
ACCESSION BX828572
VERSION BX828572.1 GI:42462677
KEYWORDS HTC; GSLT cDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1654)
Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,M., Quetier,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
REFERENCE Unpublished
AUTHORS 2 (bases 1 to 1654)
Genoscope.
TITLE Direct Submission
```

```
JOURNAL Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY Cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)
COMMENT - Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
length
http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis.
FEATURES
source Location/Qualifiers
1..1654
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="GSLTSL182D01"
/tissue_type="Silique"
/ecotype="Col-0"
/plasmid="pCMVSPORT_6"
complement(1..1654)
/gene="At4g38240"
ORIGIN
Query Match 41.6%; Score 722.4; DB 4; Length 1654;
Best Local Similarity 71.4%; Pred. No. 8.4e-19;
Matches 979; Conservative 0; Mismatches 391; Indels 2; Gaps 2;
Qy 139 AAGTTTTCCTGTTTCCGGTACCTCTCTATCTTGGCTGCTGCTCGCTTCATCTACACA 198
Db 25 AGGATCTCGTGAGCTTGAGATTTCTTCTCATCCGGCAGCTTCATGTTTATCTACATC 84
Qy 199 CAGATGCGGCTTTTTCGACACAGTCAGAAATATGCAGATTCGCTTTCGCTGCAATGAA 258
Db 85 CAGATGAGGCTTTCCAGACGCAATCAGATATGCAGATCGCTCAGTTCGCTATCGAA 144
Qy 259 CGAGAAATATGTTTGAAGCAGACAGATGCTTATTTGACCAATAGCTTGCAGCA 318
Db 145 TCTGAGAACCAATGCACTAGTCAATGCGAGGCTCATAGATGAAGTTAGCATCAACAG 204
Qy 319 GGAAGATAGTTCCTTGAAGACAAATGAAGCTCAGGACCGAGAGTGCAGCAATTA 378
Db 205 TCGCGATGTTTCCCTCGAAGATATGAAGAACCCCGAGGAGCAAGACTTGTGAGCTT 264
Qy 379 AGGGCTCTGTTTCAGGATCTTGAAGTAAAGGCGCATAAAAAGTTGATCGAAATGTACAG 438
Db 265 AAGGATCTAATCCAGACGCTTTGAAAAAAGGAATAGCAAACTCACTCAAGGTGGACAG 324
Qy 439 ATGCAGTGGCTGCTGTAGTTTATGCTTCAATCGGCTGATTAATCTGGAAGAGACT 498
Db 325 ATGCCTGCTGCTGTAGTTTATGCTCAGTTCGCTGAGACTATCTTGAAGAGACT 384
Qy 499 ATTAATCCATCTTAAATACCAATATCTGTTGCTCAAAATATCTCTCTTTTCATATCC 558
Db 385 GTTAATCAGTTTAAACATATCAAACTCCCGTTGCTTCAAAATATCTCTTATATCT 444
Qy 559 CAGGATGATCATCTCTGATGTCAGGAGCTTGTCTTGTAGCTATGATCAGCTGACGTAT 618
Db 445 CAGGATGATCTGATCAAGCTGTCAAGAGCAAGTCAATGAGCTAATAATCAATATAT 504
Qy 619 ATGCAGCACTCGATTTTGAACCTGTGCATATCTGAAGACCGAGGAGCTGATTCATAC 678
Db 505 ATGCAGCACTCGATTTTGAACCTGTGTCTGCTGAGTGTCTGAGTGTGAGTGTGAGT 563
Qy 679 TACAAAATGCAAGCTCATTTACAGTGGGCA-TTGGATCAGCTGTTTACAGCATATTT 737
Db 564 TACAAGATTGCACGTCCTACTACAAGTGGGCACTGGGACCAAGTTGTTTACAAACACAAAT 623
```

Qy	738	TAGCCGTGTTATCATCTAGAAAGATGATATGAAAATTGCCCGCTGATTTTTTTTGACTTTTT	797
Db	624	TAGTCGAGTGATTTATCTAGAAAGATGATATGAAAATTGCTCCAGACTTCTTTGATTACTT	683
Qy	798	TGAGCGTGGAGCTACTCTCTTGACAGAGACAAGTCGATTAATGCTATTTCTTCTTCGAA	857
Db	684	TGAGGCTGCGAGCTAGTCTCATGGATAGGATATAAACCAATTATGGCTGCTCATCATGAA	743
Qy	858	TGACAAATGGACAAATGCAGTTTGTCCAGATATCCCTATTGCTCTTTACCGCTCAGATTTTTT	917
Db	744	TGATAATGGACAGAAGCAGTTTGTGCATGATCCCTATGCGCTATACCGATCAGATTTTTT	803
Qy	918	TCCCGGCTCTGGATGGATGCTTTCAAAATCTACTTTGGGACGAATATCTCCAAAGTGCC	977
Db	804	TCCTGGCCTTTGGGTGGATGCTCAAGAGATCGTCTTTGGGATGAGTTATCACCCAAAGTGCC	863
Qy	978	AAAGGCTTACTGGGACGACTGGCTTAAGACTCAAGAGAAATCACAGAGGTGCACAAATTTAT	1037
Db	864	AAAGGCTTACTGGGATGATTTGGCTTGAGACTAAAGAGAAACCATAAAGCGCGCCAAATTCAT	923
Qy	1038	TCGCCCAGAAAGTTTGCGAACATATAATTTTGGTGAGCATGGTCTTAGTTTGGGCGAGTT	1097
Db	924	TCGACCGGAGTCTGTAGACATACAAATTTGGTGAACATGGGTCTAGTTTGGTGAGTT	983
Qy	1098	TTTCAAGCAGTATCTTGAGGCAATTAACATAAATGATGTCCAGGTTGATTGGGAAGTCAAT	1157
Db	984	TTTCAGTCAGTATCTGGATCCTATAAGTTATACGATGTACGGTGTGATTGGAATCAAA	1043
Qy	1158	GGACCTTAGTTACCTTTTGGGAGCAATTACGTGAACACTTTGGTGACTTGGTTAAAA	1217
Db	1044	GGACCTGGGATCTTGACAGAGGTTACTTTACCAAGTACTTTTCTGGCTTAGTGAGATA	1103
Qy	1218	GGCTTAAGCCCATCCATGGAGCTGATGCTGCTTCAAAAGCATTTAAACATAGATGGTGATGT	1277
Db	1104	AGCAGCACTAATTCAGGTTCTGACCTTGCTTTAAAGGCTCAAAACATTAAGTATGATGT	1163
Qy	1278	CGGTATTTCAGPACAGAGATCAACTAGACTTTTGAAAAATATCGACGGCAATTTGGCATTTTT	1337
Db	1164	TCCTTATCCGGTATAAAGACCAAGTAGAGTTTGAACGCCATTCGACGGGAAATTTTGATATT	1223
Qy	1338	TGAAGATGGNAGATGGTGATACACGTGCGAGCATATAAAGGAATAGTAGTTTTCGGTA	1397
Db	1224	TGAAGATTTGAAGGATGGTGGCTCTTAACAGCATATATTGGAGTAGTGGTGTTTTGAAT	1283
Qy	1398	CCAAAACGTCCAGACGTGATTTCCCTTGTTGGCCATGATTTCCGCTTCAACAACTCGGAATTTGA	1457
Db	1284	CCAGATAACAAGCGTGATTTCTTGGTTGGGCCAGATCTGTATTTGCTGTGGAAATCG	1343
Qy	1458	AGATACTTAAACAAAGATATGATTCGAGAGCGCGGCAAAATTTTTGACTTA	1509
Db	1344	AAATTCCTGATGCAAAACATGACTCTTCTCCAAACGAGCATATTTTTCATCTTA	1395

RESULT 4	768 bp	mRNA	linear	EST 07-MAR-2003
BG594788				
LOCUS				
DEFINITION	BG594788	EST493466	cSTS Solanum tuberosum	cdNA clone cSTS8020 5' sequence, mRNA sequence.

FEATURES  
source

```

seq primer: nmr_k.
location/Qualifiers
  1. .768
    /organism="Solanum tuberosum"
    /mol_type="mRNA"
    /cultivar="Kennebec"
    /db_xref="taxon:4113"
    /clone="CS18020"
    /tissue_type="sprouting eyes from tubers"
    /dev_stage="12-14 weeks post harvest"
    /lab_host="SOLR"
    /clone_lib="cSTS"
    /note="Vector: pBluescript SK(-); Site 1:
    xhoI; Various sizes of sprouting eyes (2mm)
    taken from tubers. The tubers were incubated
    dark for 2-3 weeks prior to sprouting. The
    frozen in liquid nitrogen immediately upon
    tubers."

```

## ORIGIN

Query Match	36.3%;	Score 630;	DB 2;	Length 768;
Best Local Similarity	94.2%;	Pred. No. 1.8e-168;		
Matches 654; Conservative	0;	Mismatches 40;	Indels	0; Gaps 0;
Qy	126	GATGAGAGGAA	CAAGTTTGTGCTGATTTCCGGTACCTCCTCATCTTGCGTCTGCTGTCG	185
Db	73	GATGAGAGGAA	CAAGTTTGTGCTGATTTTACGGTACCTTCTCGTGTGGTCTCTCG	132
Qy	186	CTTTCATCTACACAGATGCGGCTTTTTCGGACACAGTCAGATAATGCGCTTGC	245	
Db	133	CTTTCATCTACATACAGATGCGGCTTTTTCGGACACACAGTCAGATAATGACGCGCTTGC	192	
Qy	246	TGCTGCAATTTGAAGCAGAAAAATCATTTGTAACGCCAGACAGATGCTTATTCCACAGAT	305	
Db	193	TGCTGCAATTTGAAGCAGAAAAATCATTTGTAAGTCAGACCAAGATTTGTTATTCACAGAT	252	
Qy	306	TAGCCTTCGACGAAGGAAGATAGTTGCTCTCTTGAAGAACAAATGAAGCGTCAGACACAGGA	365	
Db	253	TAGCCAGCAGCAAGGAAGAGTAGTAGCTCTCTTGAAGAACAAATGAAGCGTCAGACACAGGA	312	
Qy	366	GTGCGACAAATTAAGGCTCTGTTTCAGGATCTTGAAGTGAAGGGCATAAAAAGTTGAT	425	
Db	313	GTGCCGGCAATTAAGGGCTCTTGTTCAGGATCTTGAAGTGAAGGGCATAAAAAGTTAAT	372	
Qy	426	CGGAAATGTACAGATGCCAGTGGCTGCTGTAGTTGTTATGGCTTGCAAATCGGGCTGATTA	485	
Db	373	CGGAGATGTGCAGATGCCAGTGGCAGCTGTAGTTGTTATGGCTTGCAGTCGTGCTGACTA	432	
Qy	486	CTTGGAAGAAGACTATTAAATCCATCTTAAATACCAAAATATCTGTTGGCTGCAAAATATCC	545	
Db	433	CTTGGAAGAGACTATTAAATCCATCTTAAATACCAAAATATCTGTTGCATCAAAATATCC	492	
Qy	546	TCCTTTTCATATCCAGATGGATCACATCCTGATGTTCAGGAAGCTTGCTTTGAGCTATGA	605	
Db	493	TCCTTTTCATATCCAGATGGATCAAAATCCTGATGTGAAGAAAGCTTGTGAGCTATGA	552	
Qy	606	TCAGCTGACGTATATGCAGCATTTGGATTTTGAACCTGTGTCATATCTGAAAGAACAAGGGA	665	
Db	553	TCAGCTGACGTATATGCAGCATTTGGATTTTGAACCTGTGTCATATCTGAAAGAACAAGGGA	612	
Qy	666	GCTGATTGCACTACTACAAATTCGACTCATTAACAAGTGGGCATTCGATCAGCTGTTTGA	725	
Db	613	ACTGGTTGCACTACTACAAGATTCGACGTCAATACAAGTGGGCATTTGGATCAGCTGTTTCA	672	
Qy	726	CAAGCATAAATTTTTCGCGTGTATATCATATCTAGAAAGATGATATCGAAATTTGCCCTGATTT	785	
Db	673	CAAGCATAAATTTTTCGCGTGTATATCATATCTAGAAAGATGATATCGAAATTTGCTGCTGATTT	732	

```

786 TTTTGACTTTTTTGAGGCTGAGCTACTCTCTT 819
|||||
733 TTTTGACTATTTTGAGGCTGAGCTACTCTCTT 766

BQ505948          772 bp  mRNA      linear      EST 07-MAR-2003
ESF613363 Generation of a set of potato cDNA clones for microarray
analyses mixed potato tissues Solanum tuberosum cDNA clone STMGH33
5' end, mRNA sequence.
BQ505948
BQ505948.2 GI:21921855
EST.
Solanum tuberosum (potato)
Solanum tuberosum
Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum.
1 (bases 1 to 772)
Buell,C.R., Hart,A., Baker,B., Tanksley,S., Fry,W., Smart,C.,
Restrepo,S., Griffiths,H., van der Hoeven,R., Tsai,J. and
Karamycheva,S.A.
Generation of a set of potato cDNA clones for microarray analyses
Unpublished (2002)
On Jun 10, 2002 this sequence version replaced gi:21364817.
Other_ESTs: EST613364
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@igr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: T3.
Location/Qualifiers
1..772
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec or Bintje"
/db_xref="taxon:4113"
/clone="STMGH33"
/tissue_type="mixed tissues"
/lab_host="SOLR"
/clone_lib="Generation of a set of potato cDNA clones for
microarray analyses mixed potato tissues"
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: Combination of untreated and Phytophthora
infestans-treated libraries of stolons, leaves, leaflets,
axillary buds of stem explants, petioles, germinating
eyes, tubers, or roots."

ORIGIN
Query Match          34.5%; Score 599.6; DB 5; Length 772;
Best Local Similarity 93.9%; Pred. No. 9.3e-160;
Matches 657; Conservative 0; Mismatches 39; Indels 4; Gaps 3;

QY 126 GATGAGAGGAAACAAGTTTGTGTAATTCGGTACCTTCCTCATCTGGCTGCTGTCGC 185
DB 73 GATGAGAGGAAACAAGTTTGTGTAATTCGGTACCTTCCTCATCTGGCTGCTGTCGC 132

QY 186 CTTTCATCTACACAGATCGGCTTTTGGACAGATGAGATATGCAGATCGCTTGC 245
DB 133 CTTTCATCTACAGATCGGCTTTTGGACAGATGAGATATGATAGCCGCTTGC 192

QY 246 TGTGCAATTGAAGCAGAAAATCATTTGTAAGCCAGACAGATTCCTTATTGACAGAT 305
DB 193 TGTGCAATTGAAGCAGAAAATCATTTGTAAGCCAGACAGATTCCTTATTGACAGAT 252

QY 306 TAGCCTTCAGCAGGGAAGATAGTTGCTCTTTGAAGAACAAATGAAGCGTCAGGACGGA 365
DB 253 TAGCCAGCAGGGAAGGAGTAGTAGCTCTTTGAAGAACAAATGAAGCGTCAGGACGGA 312

366 GTGCCGCAAAATTAAGGCTCTTGTTCAGGATCTTTGAAAGTAAGGGCATAAAAAGTTGAT 425
313 GTGCCGCAAAATTAAGGCTCTTGTTCAGGATCTTTGAAAGTAAGGGCATAAAAAGTTAAT 372

QY 426 CGGAAATGTACAGATGCCAGTGGCTGTAGTTGTTATGCTTGGCAATCGGCTGATTA 485
DB 373 CGGAGATGTGAGATGCCAGTGGCAGCTGTAGTTGTTATGCTTGGCTGAGTCTGCTGACTA 432

QY 486 CCTGAAAAAGACTATTAAATCCATCTTTAAATACCAATATCTCTTGGTCAAAATATCC 545
DB 433 CCTGAGAGGAGACTATTAAATCCATCTTTAAATACCAATATCTCTTGCATCAAAATATCC 492

QY 546 TCTTTTCATATCCAGGATGCATCATCTCTGATGTCAGGAGCTTGGTTTGAGCTATGA 605
DB 493 TCTTTTCATATCCAGGATGCATCAAAATCCCTGATGTAAGAAAGCTTGTCTTTGAGCTATGA 552

QY 606 TCAGCTGACGTATATGACGACCTTGGATTTTGAACCTGTGTCATCTGAAAGACCA-GGGG 664
DB 553 TCAGCTGACGTATATGACGACCTTGGATTTTGAACCTGTGTCATCTGAAAGACCA-GGGG 612

QY 665 AGCTGATTGCATCTACAAAATTCGACGTCAATTCACAAAGTGGG--CATTTGGATCAGCTGTT 722
DB 613 AACTGTTTGCATCTACAAAGATTCGACGTCAATTCACAAAGTGGGCACTTTGGATCAGCTGTT 672

QY 723 TTACAAGCATAA-TTTTAGCGGTGTTATCATCTACTAGTAGATGATATGAAATTTGCCCTG 781
DB 673 TCACAAGCATAAATTTTAGCGGTGTTATCATCTACTAGTAGATGATATGAAATTTGCTGCTG 732

QY 782 ATTTTGTGACTTTTGGAGCTGGAGCTACTCTCTTTGA 821
DB 733 ATTTTGTGACTATTTTGGAGCTGGAGCTACTCTCTTTGA 772

RESULT 6
LOCUS B1179808
DEFINITION EST520753 cSTE Solanum tuberosum cDNA clone cSTE20K20 5' sequence,
mRNA sequence.
ACCESSION B1179808
VERSION B1179808.1 GI:14645619
KEYWORDS EST.
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum.
1 (bases 1 to 724)
van der Hoeven,R., Bezzerides,J., Bachem,C., Visser,R., Cho,J.,
Chiemingo,A., Bougri,O., Buell,C.R., Ronning,C., Tanksley,S. and
Baker,B.
Generation of ESTs from in vitro grown microtubers
Unpublished (2001)
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@igr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: M13F-R.
Location/Qualifiers
1..724
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Bintje"
/db_xref="taxon:4113"
/clone="cSTE20K20"
/tissue_type="axillary buds of stem explants; growing
sink-tubers"
/dev_stage="7, 8 and 10 days"
/lab_host="SOLR"
/clone_lib="cSTE"
/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:

```

XhoI; Tissue supplied by Christian Bachem and Richard Visser (Department of Plant Breeding, Wageningen University, The Netherlands). The cSTA libraries will attempt to capture the induction and initiation/initial growth of the tuber in an in vitro system as described in Bachem et al. (Plant Journal, 1996). Small microtubers develop from axillary buds attached to stem explants when placed on a high sucrose medium (10%). Visible morphological changes occur synchronously at day five in the axillary buds. The first library, cSTA (1-20) consists of axillary buds harvested on days 1-3. This targets those genes involved in induction of the microtubers. The following libraries, cSTA (21-40) and cSTA (41-60), capture genes involved in tuber initiation and outgrowth. This library is noted as P3 in Tanksley lab notebooks."

## ORIGIN

Query Match 34.1%; Score 593; DB 2; Length 724;  
Best Local Similarity 93.9%; Pred. No. 7.1e-158; Indels 0; Gaps 0;  
Matches 617; Conservative 0; Mismatches 40;

QY 126 GATGAGGGGAACAGTTTTCGTGTGATTTCCGGTACCTCTCTATCTTGGTGTGCTGCGC 185  
DB 68 GATGAGGGGAACAGTTTTCGTGTGATTTACGGTACCTCTCTGCTGGCTGCTGCTGCGC 127

QY 186 CTTTCATCTACACAGATGCGGCTTTTTCGACACAGTCAGAAATGACAGATCGCCTTGC 245  
DB 128 CTTTCATCTACACAGATGCGGCTTTTTCGACACAGTCAGAAATGATGACCGCCTTGC 187

QY 246 TGCTGCAATTGAAGCAGAAATCATTTGACAGCCAGACAGATGCTTTATTGACACAGAT 305  
DB 188 TGCTGCAATTGAAGCAGAAATCATTTGACAGTCAGACAGATGCTTTATTGACAGAT 247

QY 306 TAGCCTGCAGCAAGGAATAGTTGCTTCTTGAAGAACAAATGAAGCGTCAGGACCCAGGA 365  
DB 248 TAGCCAGCAGCAAGGAGTAGTAGCTCTTGAAGACAAATGAAGCATCAGGACCCAGGA 307

QY 366 GTGCCGCAATTAAGGGCTCTTTTCAGGATCTTTGAAGTAAGGCGATATAAAAGTTGAT 425  
DB 308 GTGCCGCAATTAAGGGCTCTTTTCAGGATCTTTGAAGTAAGGCGATATAAAAGTTAAT 367

QY 426 CGGAAATCTACAGATGCCAGTGGCTGTGATGTTTATGGCTTGCATCGGCTGATTA 485  
DB 368 CGGAGATGTGAGATGCCAGTGGCAGCTGTAGTTGTTATGGCTTGCATCGCTACTGACTA 427

QY 486 CTTGAAAGACTATTAATCCATCTTAAATATCAAAATATCTGTGCGTCAAAATATCC 545  
DB 428 CTTGAGAGGACTATTAATCCATCTTAAATATCAAAATATCTGTGATCAAAATATCC 487

QY 546 TCTTTTCATATCCAGGATGGATCATCTCCTGTGATGTCAAGAGCTTGTCTTTGAGCTATGA 605  
DB 488 TCTTTTCATATCCAGGATGGATCAATCTCCTGTGATGTCAAGAGCTTGTCTTTGAGCTATGG 547

QY 606 TCAGCTGACGTATATGACAGCTTGGATTTTGAACCTGTGCTATCTGAAGACACAGGGA 665  
DB 548 TCAGCTGACGTATATGACAGCTTGGATTTGAACCTGTGCTATCTGAAGACACAGGGA 607

QY 666 GCTGATTCGATCTACAAATTTGACAGCTCATATTCAAGTGGGCATTGGATCAGCTGTTTTTA 725  
DB 608 ACTGGTGCATCTACAGATTTGACAGCTCATATTCAAGTGGGCATTGGATCAGCTGTTTTCA 667

QY 726 CAAGCATAAATTTAGCCGTGTTATCATATCTAGAGATGATATGGAATTTGCCCTCTGA 782  
DB 668 CAAGCATAAATTTAGCCGTGTTATCATATCTAGAGATGATATGGAATTTGCTGCTGA 724

RESULT 7  
BG592816  
LOCUS  
DEFINITION  
EST491494 cSTS Solanum tuberosum cDNA clone cSTS2B9 5' sequence,  
mRNA sequence.  
ACCESSION  
BG592816  
VERSION  
BG592816.1 GI:13610956

KEYWORDS  
SOURCE

## ORGANISM

Solanum tuberosum (potato)

Solanum tuberosum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
asterids; lamids; Solanales; Solanaceae; Solanum.

## REFERENCE

1 (bases 1 to 642)

## AUTHORS

van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Chiemingo,A.,

Bougri,O., Buell,C.R., Ronning,C., Tanksley,S. and Baker,B.

## TITLE

Generations of ESTs from sprouting potato eyes

## JOURNAL

Unpublished (2000)

## COMMENT

Contact: Robin Buell  
The Institute for Genomic Research  
9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@tigr.org

This clone can be obtained from the University of Arizona Genomics  
Institute. Orders can be made through URL:

http://genome.arizona.edu/orders/

Seq primer: M13F-R.

## FEATURES

## Location/Qualifiers

1..642

/organism="Solanum tuberosum"

/mol\_type="mRNA"

/cultivar="Kennebec"

/db\_xref="taxon:4113"

/clone="cSTS2B9"

/tissue\_type="sprouting eyes from tubers"

/dev\_stage="12-14 weeks post harvest"

/lab\_host="SOLR"

/clone\_lib="cSTS"

/note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:

XhoI; Various sizes of sprouting eyes (2mm to 15mm) were

dark for 2-3 weeks prior to sprouting. The eyes were

frozen in liquid nitrogen immediately upon removal from

tubers."

## ORIGIN

Query Match 33.4%; Score 580.2; DB 2; Length 642;  
Best Local Similarity 94.1%; Pred. No. 3.1e-154;  
Matches 603; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 420 GTTGTATCGGAAATGTACAGATGCCAGTGGCTGTAGTTGTTATGGCTTGCATCGGC 479  
DB 2 GTTAAATCGGAGATGTGCAGATGCCAGTGGCAGCTGTAGTTGTTATGGCTTGCATCGTCG 61

QY 480 TGATTACCTCGGAAAGACTATTAAATCCATCTTAAATACCAATATCTGTGCGTCAAA 539  
DB 62 TGACTACCTGGAGAGACTATTAAATCCATCTTAAATACCAATATCTGTGATCAAA 121

QY 540 ATATCCTCTTTTCATATCCAGGATGGATCATCTCTGTATGTTCAGGAAGCTTGTCTTTGAG 599  
DB 122 ATATCCTCTTTTCATATCCAGGATGGATCAATCTCTGTATGTAAAGAAAGCTTGTCTTTGAG 181

QY 600 CTATGATCAGCTCACTATATGACAGCTTGGATTTTGAACCTGTGCATCTGAAGACC 659  
DB 182 CTATGATCAGCTCACTATATGACAGCTTGGATTTTGAACCTGTGCATCTGAAGACC 241

QY 660 AGGGGAGCTGATTGCTATCTACAAATTTGCACGCTCATTTCAAGTGGGCATTGGATCAGCT 719  
DB 242 AGGGGAACCTGGTTGCTATCTACAGATTGCACGCTCATTTCAAGTGGGCATTGGATCAGCT 301

QY 720 GTTTTACAAGCATAAATTTTAGCCGTGTTATCATATCTAGAAAGATGATATGGAATTTGCCCC 779  
DB 302 GTTTTACAAGCATAAATTTTAGCCGTGTTATCATATCTAGAAAGATGATATGGAATTTGCTGC 361

QY 780 TGATTTTTCGACTTTTTCGAGCTGGAGCTACTCTTCTTGACAGAGACAAGTCGATTAT 839  
DB 362 TGATTTTTCGACTTTTTCGAGCTGGAGCTACTCTTCTTGACAGAGACAAGTCGATTAT 421

QY 840 GGCTATTTCCTTCTTGAAGTACAAATGGACAAATATGCAAGTTTGTCCAAAGATCCTTATGCTCT 899  
DB 422 GGCTATTTCCTTCTTGAAGTACAAATGGACAAAGGAGTTCGTCCAAAGATCCTGATGCTCT 481

```

Qy 900 TTACGGCTCAGATTTTCCCGGCTTGGATGATGCTTTCAAAATCTACTTGGGACGA 959
Db 482 TTACCGGTCAGACTTTTCTCGTCTGTGGATGATGCTTTCAAAATCAACTTGGTCCGA 541
Qy 960 ATTATCTCCAAAGTGGCCAAAGGCTTACTGGGACGACTGGCTAAGACTCAAGAGAATCA 1019
Db 542 ACTATCTTCAAGTGGCCAAAGGCTTACTGGGATGACTGGCTAAGGCTGAAGAAATCA 601
Qy 1020 CAGAGGTCGACAAATTTATTCGCCCCAGAAAGTTTGAGAACAT 1060
Db 602 CAGAGGTCGACAAATTTATTCACCCAGAAAGTTTGCAAAACGT 642

RESULT 8
LOCUS AI486771 643 bp mRNA linear EST 18-MAY-2001
DEFINITION EST245093 tomato ovary, TAMU Lycopersicon esculentum cDNA clone
cLED11D20, mRNA sequence.
ACCESSION AI486771 GI:4382142
VERSION AI486771
KEYWORDS Lycopersicon esculentum (Solanum lycopersicum)
SOURCE Lycopersicon esculentum
ORGANISM Lycopersicon esculentum
REFERENCE 1 (bases 1 to 643)
AUTHORS Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T.,
Holt,I.E., Liang,F., Upton,J., Ronning,C.M., Craven,M.B.,
Fujii,C.Y., Bowman,C.L., Nierman,W., Fraser,C.M., Venter,J.C.,
Martin,G.B., Tanksley,S.D. and Giovannoni,J.
Generation of ESTs from tomato carpel tissue
Contact: CUGI
Unpublished (1999)
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html.
Location/Qualifiers
1..643
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLED11D20"
/tissue_type="carpel"
/dev_stage="5 days pre-anthesis to 5 days post-anthesis"
/lab_host="XL1-Blue MRF"
/clone_lib="tomato ovary, TAMU"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; cLED - Tomato Carpel EST Library. Oligodt-primed and
directionally cloned cDNA in vector Lambda ZAP II with 5'
and 3' ends located at the EcoRI and XhoI sites,
respectively."

ORIGIN
Query Match 33.3%; Score 578.6; DB 1; Length 643;
Best Local Similarity 93.9%; Pred. No. 9e-154;
Matches 602; Conservative 0; Mismatches 39; Indels 0; Gaps 0;

Qy 420 GTTGATCGGAATGTACAGATGCCAGTGGCTGCTGTAGTTGTATGGCTTGAATCGGC 479
Db 1 GTTAATCGGAATGTGCAGATGCCAGTGGCAGCTGTAGTTGTATGGCTTGCAGTCGTTTC 60
Qy 480 TCATTACCTGGAAAGACATTAATAATCCATCTTAAATACCAATATCTGTTGGCGTCAA 539
Db 61 TGACTACCTGGGAAGACATTAATAATCCATCTTAAATACCAACATCTGTTGCATCAA 120
Qy 540 ATATCTCTTTTCATATCCAGGATGGATCAGATCCTGTATGTGAGAGCTTGTCTTGAG 599
Db 121 ATATCTCTTTTCATATCCAGGATGGATCAGATCCTGTATGTGAGAGCTTGTCTTGAG 180

```

```

Qy 600 CTATGATCAGCTCAGCTATATGACGACATTTGGATTTTGAACCTGTGATCTGAAGACC 659
Db 181 CTATGATCAACTGACCTATATGACGACATTTGGATTTTGAACCTGTGATCTGAAGACC 240
Qy 660 AGGGGAGCTGATTCATACACTACAAAATTCGACGCTCATTAAGTGGGCAATGGATCAGCT 719
Db 241 AGGGGAACTGGTTCATACACTACAAAATTCGACGCTCATTAAGTGGGCAATGGATCAGCT 300
Qy 720 GTTTTCAAGCATAAATTTTAGCCGCTGTATCATACTAGAAGATGATATGGAATTTGCCCC 779
Db 301 GTTTTCAAGCATAAATTTTAGCCGCTGTATCATACTAGAAGATGATATGGAATTTGCCGC 360
Qy 780 TGAATTTTGTGACTTTTTCGAGGCTGGAGCTACTCTCTTTCGACAGACAAGTCGATAT 839
Db 361 TGAATTTTGTGACTTTTTCGAGGCTGGAGCTACTCTCTTTCGACAGACAAGTCGATAT 420
Qy 840 GGCTATTTCTTCTTGGATGACATGACAAATTCAGTTTGTGTCACAGATCCTTATGCTCT 899
Db 421 GGCTATTTCTTCTTGGATGACATGACAAATTCAGTTTGTGTCACAGATCCTTATGCTCT 480
Qy 900 TTACCGCTCAGATTTTTCGCCGCTTGGATGATGCTTTCAAAATCTACTTGGGACGA 959
Db 481 TTACCGCTCAGACTTTTTCGCCGCTTGGATGATGCTTTCAAAATCAACTTGGTCCGA 540
Qy 960 ATTATCTCCAAAGTGGCCAAAGGCTTACTGGGACGACTGGCTAAGACTCAAGAGAATCA 1019
Db 541 ACTATCTCCAAAGTGGCCAAAGGCTTACTGGGATGACTGGCTGAGGCTCAAGAAAAATCA 600
Qy 1020 CAGAGGTCGACAAATTTATTCGCCCCAGAAAGTTTGAGAACAT 1060
Db 601 CAGAGGTCGACAAATTTATTCGCCCCAGAAAGTTTGAGAACAT 641

RESULT 9
LOCUS BG889872 661 bp mRNA linear EST 07-MAR-2003
DEFINITION EST515723 cSTD Solanum tuberosum cDNA clone cSTD15H10 5' sequence,
mRNA sequence.
ACCESSION BG889872 GI:14266958
VERSION BG889872.1
KEYWORDS Solanum tuberosum (potato)
SOURCE Solanum tuberosum
ORGANISM Solanum tuberosum
REFERENCE 1 (bases 1 to 661)
AUTHORS van der Hoeven,R., Bezzerides,J., Ewing,E., Cho,J., Chiemiango,A.,
Bougrl,O., Buell,C.R., Ronning,C., Tanksley,S. and Baker,B.
Generations of ESTs from dormant potato tubers
Unpublished (2001)
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@igr.org
This clone can be obtained from the University of Arizona Genomics
Institute. Orders can be made through URL:
http://genome.arizona.edu/orders/
Seq primer: M13F-R.
Location/Qualifiers
1..661
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cSTD15H10"
/tissue_type="dormant tuber"
/dev_stage="one month post-harvest"
/lab_host="SOLR"
/clone_lib="cSTD"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; This library targets genes expressed in dormant
tubers. This library was made from sections of dormant

```

tuber, avoiding the buds and epidermis. Tubers were stored for one month post-harvest at 40C. The tuber was peeled, well away from the surface. Then it was chopped into 1-2 mm cubes and immediately frozen in liquid nitrogen. This library is noted as P4 in Tanksley lab notebooks."

## ORIGIN

Query Match 31.9%; Score 553.6; DB 2; Length 661;  
 Best Local Similarity 94.4%; Pred. No. 1.3e-146;  
 Matches 574; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

Qy 126 GATGAGAGGGAACAAGTTTGTCTGTGATTTCCTCGGTACCTCTCTCATCTTTGGCTGTGCTGCG 185  
 Db |||||||  
 Qy 54 GATGAGAGGGAACAAGTTTGTCTGTGATTTCCTCGGTACCTCTCTCATCTTTGGCTGTGCTGCG 113  
 Db |||||||

Qy 186 CTTTCATCTACACACAGATGCGGCTTTTGGGACACAGTCAGAAATATGACAGATCGCCCTTGC 245  
 Db |||||||

Qy 114 CTTTCATCTACACACAGATGCGGCTTTTGGGACACAGTCAGAAATATGACAGATCGCCCTTGC 173  
 Db |||||||

Qy 246 TGCTGCAATTGAACGAGAAATCATTTGACAGCCAGACAGATGCTTTATTGACCCAGAT 305  
 Db |||||||

Qy 174 TGCTGCAATTGAACGAGAAATCATTTGACAGCCAGATGCTTTATTGACCAAGAT 233  
 Db |||||||

Qy 306 TAGCCTGACGACGAAAGATAGTTGCTTTTGAAGAACAAATGAAGCGTCAGGACCAAGGA 365  
 Db |||||||

Qy 234 TAGCAGCAGCAGAGAGATAGTAGCTCTTGAAGAACAAATGAAGCGTCAGGACCAAGGA 293  
 Db |||||||

Qy 366 GTGCCGCAATTAAGGCTCTTTGTCAGGATCTTTGAAAGTAAGGGCATAAAAAAGTTGAT 425  
 Db |||||||

Qy 294 GTGCCGCAATTAAGGCTCTTTGTCAGGATCTTTGAAAGTAAGGGCATAAAAAAGTTGAT 353  
 Db |||||||

Qy 426 CGGAATGTACAGATGCCAGTGGCTGTGTAGTTGTTATGGCTTGGCAATCGGCTGATTA 485  
 Db |||||||

Qy 354 CGSAGATGTGAGATGCCAGTGGCAGCTGTAGTTGTTATGGCTTGGCAATCGGCTGATTA 413  
 Db |||||||

Qy 486 CTTGGAAGAAAGACTATTAATCCATCTTAAATAACCAATATCTCTGCGTCAAAATATCC 545  
 Db |||||||

Qy 414 CTTGAGAGGACTATTAATCCATCTTAAATAACCAATATCTCTGCGTCAAAATATCC 473  
 Db |||||||

Qy 546 TCTTTTCATATCCAGGATGGATCACATCTCTGATGTGAGGAAGCTTCTTTTGAGCTATGA 605  
 Db |||||||

Qy 474 TCTTTTCATATCCAGGATGGATCAAATCTCTGATGTAAAGAAAGCTCTCTTTGAGCTATGA 533  
 Db |||||||

Qy 606 TCAGCTACGCTATATGAGCACTTGGATTTTGAACCTGTGCATCTGAAAGACCAAGGGA 665  
 Db |||||||

Qy 534 TCAGCTACGCTATATGAGCACTTGGATTTTGAACCTGTGCATCTGAAAGACCAAGGGA 593  
 Db |||||||

Qy 666 GCTGATTGCATCTACAAAATTGACGCTCATTAACAAGTGGGCATTTGGATCAGCTGTTTTA 725  
 Db |||||||

Qy 594 ACTGGTTGCATCTACAAAGATTGACGCTCATTAACAAGTGGGCATTTGGATCAGCTGTTTTA 653  
 Db |||||||

Qy 726 CAAGCATA 733  
 Db |||||||

Qy 654 CAAGCATA 661  
 Db |||||||

## RESULT 10

AW979500  
 LOCUS  
 DEFINITION EST310521 tomato root deficiency, Cornell University Lycopersicon  
 esculentum cDNA clone CLEW209 5', mRNA sequence.

AW979500  
 ACCESSION

AW979500  
 VERSION

AW979500.1 GI:8171019  
 KEYWORDS

EST.  
 SOURCE

Lycopersicon esculentum (Solanum lycopersicum)

## ORGANISM

Lycopersicon esculentum  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.  
 1 (bases 1 to 673)  
 van der Hoeven,R.S., Garvin,D.F., Matern,A.L., Holt,I.E., Liang,F.,  
 Upton,J., Hansen,T.S., Ronning,C.M., Craven,M.B., Bowman,C.L.,  
 Nierman,W., Fraser,C.M., Venter,J.C., Martin,G.B., Giovannoni,J.J.

## REFERENCE

## AUTHORS

and Tanksley,S.D.  
 Generation of ESTs from tomato nutrient-deficient roots  
 Unpublished (1999)

## TITLE

## JOURNAL

## COMMENT

Contact: CUGI  
 Clemson University Genomics Institute  
 Clemson University  
 100 Jordan Hall, Clemson, SC 29634, USA  
 Email: <http://www.genome.clemson.edu/orders/index.html>  
 5 prime sequence.

## FEATURES

## source

1..673  
 Location/Qualifiers  
 /organism="Lycopersicon esculentum"  
 /mol\_type="mRNA"  
 /cultivar="TA496"  
 /db\_xref="taxon:4081"  
 /clone="CLEW209"  
 /tissue\_type="roots"  
 /dev\_stage="5-6 weeks old"  
 /clone\_lib="tomato root deficiency, Cornell University"  
 /notes="Vector: pBluescript SK-; Site\_1: 5' EcoRI; Site\_2:  
 3' XhoI; supplier: Tanksley; Tissue supplied by Dave  
 Garvin (USDA-ARS, Ithaca, NY 14850). Roots were harvested  
 from plants grown under the following  
 deficiencies/stresses: 10 mM Al, Zn, P, K, Fe, N), and  
 mRNA was isolated from individual treatments. Proportional  
 aliquots of mRNA of each treatment were mixed and used for  
 library construction."

## ORIGIN

Query Match 31.0%; Score 537.8; DB 1; Length 673;  
 Best Local Similarity 93.1%; Pred. No. 4.4e-142;  
 Matches 563; Conservative 0; Mismatches 42; Indels 0; Gaps 0;

Qy 126 GATGAGAGGGAACAAGTTTGTCTGTGATTTCCTCGGTACCTCTCTCATCTTTGGCTGTGCTGCG 185  
 Db |||||||

Qy 69 GATGAGAGGGAACAAGTTTGTCTGTGATTTCCTCGGTACCTCTCTCGTGTGCTGCTCTCGC 128  
 Db |||||||

Qy 186 CTTTCATCTACACACAGATGCGGCTTTTGGGACACAGTCAGAAATATGACAGATCGCCTTGC 245  
 Db |||||||

Qy 129 CTTTCATCTACACAGATGCGGCTTTTGGGACACAGTCAGAAATATGACAGATCGCCTTGC 188  
 Db |||||||

Qy 246 TGCTGCAATTGAACGAGAAATCATTTGACAGCCAGACAGATGCTTTATTGACCCAGAT 305  
 Db |||||||

Qy 189 TGCTGCAATTGAACGAGAAATCATTTGACAGTCAGCTAGGTTGCTTTATTGACCAAGAT 248  
 Db |||||||

Qy 306 TAGCCTGACGACGAGGAAGATAGTCTCTTGAAGAACAAATGAAGCGTCAGGACCAAGGA 365  
 Db |||||||

Qy 249 TAGCCAGCAGCAAGGAAGAGTAGTGGCTCTTGAAGAACAAATGAAGCGCCAGGACCAAGGA 308  
 Db |||||||

Qy 366 GTGCCGCAATTAAGGCTCTTTGTCAGGATCTTTGAAAGTAAGGGCATAAAAAAGTTGAT 425  
 Db |||||||

Qy 309 GTGCCGCAATTAAGGCTCTTTGTCAGGATCTTTGAAAGTAAGGGCATAAAAAAGTTGAT 368  
 Db |||||||

Qy 426 CGGAATGTACAGATGCCAGTGGCTGTGTAGTTGTTATGGCTTGGCAATCGGCTGATTA 485  
 Db |||||||

Qy 369 CGGAATGTACAGATGCCAGTGGCTGTGTAGTTGTTATGGCTTGGCAATCGGCTGATTA 428  
 Db |||||||

Qy 486 CTTGGAAGAAAGACTATTAATCCATCTTAAATAACCAAAATATCTTTGCGTCAAAATATCC 545  
 Db |||||||

Qy 429 CTTGGAAGAAAGACTATTAATCCATCTTAAATAACCAAAATATCTTTGCGTCAAAATATCC 488  
 Db |||||||

Qy 546 TCTTTTCATATCCAGGATGGATCACATCTCTGATGTGAGGAAGCTTGTGCTTTGAGCTATGA 605  
 Db |||||||

Qy 489 TCTTTTCATATCCAGGATGGATCGAATCTCTGATGTAGAAAGCTTGTGCTTTGAGCTATGA 548  
 Db |||||||

Qy 606 TCAGCTGACGCTATATGACGACCTTGGATTTTGAACCTGTGCATCTGAAAGACCAAGGGA 665  
 Db |||||||

Qy 549 TCAGCTGACGCTATATGACGACCTTGGATTTTGAACCTGTGCATCTGAAAGACCAAGGGA 608  
 Db |||||||

Qy 666 GCTGATTGCATCTACAAAATTGACGCTCATTAACAAGTGGGCATTTGGATCAGCTGTTTTA 725  
 Db |||||||

Qy 609 ACTGGTTGCATCTACAAAGATTGCAGCTCATTAACAAGTGGGCATTTGGATCAGCTGTTCA 668  
 Db |||||||

```

Oy 726 CAAGC 730
    |||||
Db 669 CAAGC 673

RESULT 11
AW735827
LOCUS EST335827 659 bp mRNA linear EST 18-MAY-2001
DEFINITION Lycopersicon esculentum cDNA clone cfoa5111 5', mRNA sequence.
ACCESSION AW735827
VERSION AW735827.1 GI:7642666
KEYWORDS EST.
SOURCE Lycopersicon esculentum (Solanum lycopersicum)
ORGANISM Lycopersicon esculentum
REFERENCE
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 659)
van der Hoeven,R.S., Bezzerides,J.L., Matern,A.L., Holt,I.E.,
Liang,F., Hansen,T., Craven,M.B., Bowman,C.B., Romning,C.M.,
Niernan,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and
Tanksley,S.D.
Generation of ESTs from tomato flower tissue, 0-3 mm buds
Contact: CUGI
Unpublished (1999)
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html
5 prime sequence.
LOCATION/Qualifiers
1..659
/organism="Lycopersicon esculentum"
/mol_type="mRNA"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="CT0A5111"
/tissue_type="Flower"
/dev_stages="0-3mm buds"
/clone_lib="tomato flower buds 0-3 mm, Cornell University"
/notes="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
XhoI; supplier: Tanksley; Flower buds and flowers were
taken from greenhouse plants (4-8 wks old, TA496). They
were immediately frozen in liquid nitrogen and then
size-separated while remaining frozen."

FEATURES
source
ORIGIN
Query Match 30.3%; Score 527; DB 1; Length 659;
Best Local Similarity 93.2%; Pred. No. 5.3e-139;
Matches 551; Conservative 0; Mismatches 40; Indels 0; Gaps 0;

Oy 126 GATGAGAGGGAACAAGTTTGTGTTGATTTCCGGTACCTCCCTCATCTTGGCTGCTGTCG 185
    |||||
Db 69 GATGAGAGGAACAAGTTTGTGTTGATTTACGGTACCTCTCGTGGTGGCTCTCGC 128

Oy 186 CTTTCATCTACACAGATCGCGCTTTTCGACACAGTCAGATATGCAATATGCGCTTGC 245
    |||||
Db 129 CTTTCATCTACATACAGATCGCGCTTTTCGTCACACAGTCAGATATGCAATATGCGCTTGC 188

Oy 246 TCGTGAATTTGAAGCAGAAAATCAATTTGACAGCCAGACCAAGTTGCTTATTGACCAT 305
    |||||
Db 189 TCGTGGGATTTGAAGCAGAAAATCAATTTGACAGTCAGATAGTGTCTTATTGACAA 248

Oy 306 TAGCCTGACGAGGGAAGATAGTCTCTTTGAAGAACAAATGAAGCGTCAGGACGAGCA 365
    |||||
Db 249 TAGCCAGCAGCAGGAGAGTAGTGGCTCTTGAAGAACAAATGAAGCGGCGGACGAGCA 308

Oy 366 GTGCCGACAAATTAAGGGCTCTTCTTCAGGATCTTTGAAGTAAGGGCATAAAAAGTTGAT 425
    |||||
Db 309 GTGCCGACAAATTAAGGGCTCTTCTTCAGGATCTTTGAAGTAAGGGCATAAAAAGTTAAT 368
    |||||
Oy 426 CGGAATGTACAGATGCCAGTGGCTGCTGTAGTTGTTTATGGCTTGAATCGGCTGATTA 485
    |||||

```

```

Db 369 CGGAATGTGAGATGCCAGTCCAGCTAGTGTATGCTTGCAGTCTTCTGACTA 428
    |||||
Oy 486 CTTGGAAGAAGACTATTAAATCCATCTTAAATACCAATATCTGCTGCTCAAAATATCC 545
    |||||
Db 429 CTTGGAAGAAGACTATAAAATCCATCTTAAATACCAAAACATCTGTTGCATCAAAATATCC 488
    |||||
Oy 546 TCTTTTCATATCCAGGATGATCACATCTCTGATGTGAGGAAGCTTGTCTTTGAGCTATGA 605
    |||||
Db 489 TCTTTTCATATCCAGGATGATCGAATCTCTGATGTAGAAAGCTTGTCTTTGAGCTATGA 548
    |||||
Oy 606 TCAGCTGACGTATATGACGACCTTGGATTTTGAACCTGTGCATCTGAACACAGGCGGA 665
    |||||
Db 549 TCAACTGACGTATATGACGACCTTGGATTTTGAACCTGTGCATCTGAACACAGGCGGA 608
    |||||
Oy 666 GCTGATTGCATCTACAAAATTTGACCTGCTTATCAAGTGGCATTTGGATCA 716
    |||||
Db 609 ACTGTTGCATCTACAAAATTTGACCTGCTTATCAAGTGGCATTTGGATCA 659
    |||||

RESULT 12
BM779745
LOCUS EST590321 KV2 Medicago truncatula cDNA clone pKV2-52N22, mRNA
DEFINITION sequence.
ACCESSION BM779745
VERSION BM779745.1 GI:19109675
KEYWORDS EST.
SOURCE Medicago truncatula (barrel medic)
ORGANISM Medicago truncatula
REFERENCE
AUTHORS VandenBosch,K., Endre,G., Hur,J., Beremand,P., Town,C.D., Van
Aken,S., Utterback,T., Cheung,F., Tsai,J. and Fraser,C.M.
ESTs from roots of Medicago truncatula 48 hr after inoculation with
Sinorhizobium meliloti
Unpublished (2002)
Contact: VandenBosch K
Department of Plant Biology
University of Minnesota
220 Biosci Center, 1445 Gortner Ave, St. Paul, MN 55108, USA
Tel: 612 624 2755
Fax: 612 625 1738
Email: kvandenb@cbs.umn.edu
TIGR sequence name: MTACP83TK More information is available at:
www.medicago.org
Seq primer: SKmod (CTA GAA CTA gtg gat CC) .
FEATURES
source
1..843
/organism="Medicago truncatula"
/mol_type="mRNA"
/cultivar="genotype A17"
/db_xref="taxon:3880"
/clone="pKV2-52N22"
/tissue_type="Seedling roots"
/dev_stages="2 days post-inoculation with Sinorhizobium
meliloti"
/lab_host="E. coli strain SOLR"
/clone_lib="KV2"
/notes="Vector: pBluescript SK -; Site 1: EcoRI; Site 2:
XhoI; cDNA was prepared from polyA+ enriched RNA. The
cDNA was directionally ligated into the Unipak XR vector
from Stratagene and packaged using Gigapack III Gold
packaging extracts. Plasmids containing cDNA inserts
were excised from the recombinant lambda-Zap phage using
Ex-assist helper phage and propagated in SOLR cells."

```

```

ORIGIN
Query Match 29.8%; Score 517.2; DB 3; Length 843;
Best Local Similarity 75.9%; Pred. No. 3.7e-136;

```



Matches 639; Conservative 0; Mismatches 203; Indels 0; Gaps 0;			
Qy	142	TTTTGCTGTGATTTCCGGTACCTCCTCATCTTGGCTGCTGTCGCTTCATCTACACAG	201
Db	2	TTTTGGTGTGATTTCCGGTATCTTCTCTTTGTCGAGCCTTGGTGTTCATCTACATCCAG	61
Qy	202	ATGGGGCTTTTTCGACACAGCTCAGAATATGCAGATCGCCTTGCTGCTGCAATTTGAAGCA	261
Db	62	ATGAGACTTTTTCGATCACAATCACAATATGCTGATCGCCTGCTGCTGCTATTTGAAGCT	121
Qy	262	GAATATCATTGTACAAGCCAGACAGATTTGCTTATTTGACAGATTAGCTTGCGACAAGGA	321
Db	122	GAATATCATTGTACAAGCTCAAATCCGATCACTCATTTGATCAGATTAGCTTGCAACAAGGA	181
Qy	322	AGAATAGTTGCTCTTTGAAGAACAAATGAAGCGTCAGGACGAGGAGTGCAGCAATTTAAGG	381
Db	182	CGCATTTCTCGACTTTACAACAAGAAAGGACCGTCGAGAGCAAGAATGTAGCCAATCAAG	241
Qy	382	GCTCTTGTTCAGGATCTTTGAAGTAAGGGCATAAAAAGTTGATCGGAAATGTACAGATG	441
Db	242	TCTCTTGTTCGAATCTTTGAAGAAAGATGTACGAAGGCTGATTTGATTAAGTCCAGTT	301
Qy	442	CCAGTGGCTGCTGTAGTTGTTATGGCTTTGCAATCGGGCTGATTACCTGGAAAAGACTATT	501
Db	302	CCAGTGGCAGCTGTTGTCATCATGSCATGTAATCGTGTGATTACCTGGAGAGACTATC	361
Qy	502	AAATCCATCTTAAATACCAAAATATCTGTTCGCTCAAAATATCTCTTTTCATATCCAG	561
Db	362	AATTCTGTATTGAATAATATCAAGGCCCATTTCTTCAAGATTTCTCTTATTTGTATCACAG	421
Qy	562	GATGGATCACATCTGTATGTCAGGAAGCTTGCTTTGAGCTATGATCAGCTGACGTATATG	621
Db	422	GATGGATCAAAATTCAGACGCTCAAAAGAAAGCTCTGAGCTACGATGAGCTATCTCATATG	481
Qy	622	CAGCACTTGGATTTTGAACCTGTGCATACTGAAAGACCAGGGGAGCTGATTGCATATAC	681
Db	482	CAGCACTTAGATTTTGAAGCGGTTCAAACCTGAAGAGCCAGAGAGTTAATTTGCTTACTAT	541
Qy	682	AAATATGCACGTATTAACAAGTGGGCATTTGATCAGCTGTTTTCACAAGCATATAATTTAGC	741
Db	542	AAGATTGCACGTATTAACAAGTGGCTCTTGCTCAACTGTTTCGACAAGCATAACTTCAAC	601
Qy	742	CGTGTATCATACTAGAGATGATATGGAATTCGCCCTGATTTTTCCTTTTGTGACTTTTTCAG	801
Db	602	AGAGTCATCATCTTCGAAGATGACATGGAATAGCACCCTGATTTCTTTGATTAATTTTGA	661
Qy	802	GCTGGAGCTACTCTCTTTCAGACAGACAAGTCGATTTATGGCTATTTCTTTTGGAAATGAC	861
Db	662	GCTATGGCACTCTCTTTGACAAGGATAAGTCCATTAATGGCTGTTTGTCAATGGAATGAC	721
Qy	862	AATGGACAAATGCAAGTTTGTCCAAGATCTCTTATGCTCTTTACCGCTCAGATTTTTTCCC	921
Db	722	AATGGACAGAAGCAGTTTGTACATGATCCATATGAATTTTATCGCTCGGACTCTTTCTCT	781
Qy	922	GCTCTTGGATGGATCTTTCAAAATCTACTTGGACGCAATTAATCTCCAAAGTGCCCAAG	981
Db	782	GGACTAGGATGGATGGCTGGGTAGATCTACATGGGATGAACCTATCACCANAATGGCCGAA	841
Qy	982	GC 983	
Db	842	GC 843	

RESULT 13  
AW931180  
LOCUS  
DEFINITION EST357023 tomato fruit mature green, TAMU Lycopersicon esculentum  
CDNA clone cLEF43B14 5', mRNA sequence.  
AW931180  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AW931180 574 bp mRNA linear EST 18-MAY-2001  
EST357023 tomato fruit mature green, TAMU Lycopersicon esculentum  
CDNA clone cLEF43B14 5', mRNA sequence.  
AW931180  
AW931180.1 GI:8106581  
Lycopersicon esculentum (Solanum lycopersicum)  
Lycopersicon esculentum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.  
1. (bases 1 to 574)  
Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F., Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S., Ronning,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.  
Generation of ESTs from tomato fruit tissue  
Unpublished (1999)  
Contact: CUGI  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>  
5 prime sequence.  
Location/Qualifiers  
1. 574  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone="cLEF43B14"  
/tissue\_type="fruit pericarp"  
/dev\_stage="mature green (3-5 days pre-ripening)"  
/lab\_host="SOLR"  
/clone\_lib="tomato fruit mature green, TAMU"  
/notes="Vector: pBlueScript SK(-); Site.1: EcoRI; Site.2: XhoI; cLEF - Fruit were tagged at the 1cm stage and harvested 3-5 days prior to ripening. Fruit were cut in half to verify the seeds were indeed 'immature' and the seeds and locules were discarded prior to freezing the pericarp"  
ORIGIN  
Query Match 29.6%; Score 514; DB 1; Length 574;  
Best Local Similarity 93.7%; Pred. No. 2.7e-135;  
Matches 535; Conservative 0; Mismatches 36; Indels 0; Gaps 0;  
Qy 436 CAGATGCCAGTGGCTGCTGTAGTTGTTATGGCTTGCATCGGCTGATTACCTGGAAAAG 495  
Db 4 CAGATGCCAGTGGCAGCTGTAGTTGTTATGGCTTGCAGTCTGTGACTACCTGGAGAAG 63  
Qy 496 ACTATTAAATCCATCTTAAATACCAAAATATCTGTTCGTCGTCAAAATATCTCTTTTCATA 555  
Db 64 ACTATAAATCCATCTTAAATACCAAAATATCTGTTCGTCGTCAAAATATCTCTTTTCATA 123  
Qy 556 TCCAGGATGGATCACATCTGTATGTCAGGAAGCTTGTTCGTCGTCATGATGATCAGCTGACG 615  
Db 124 TCCAGGATGGATCGAATCCTGATGTTAGAAAGCTTGTTCGTCGTCATGATGATCACTGACG 183  
Qy 616 TATATGACGACCTTGGATTTTGAACCTGTGCATACTGAAAGACCAGGGGAGCTGATTGCA 675  
Db 184 TATATGACGACCTTGGATTTTGAACCTGTGCATACTGAAAGACCAGGGGAGCTGATTGCA 243  
Qy 676 TACTACAAATTCGACGTCATTAACAAGTGGGCATTTGGATCAGCTGTTTTCACAGCATAAAT 735  
Db 244 TACTACAAAGTTGCAGCTCATTAACAAGTGGGCATTTGGATCAGCTGTTTTCACAGCATAAAT 303  
Qy 736 TTTAGCCGTGTTATCATCTACTAGAGATGATATGGAATTCGCCCTGATTTTTTTTGACTTT 795  
Db 304 TTTAGCCGTGTTATCATCTACTAGAGATGATATGGAATTCGCCCTGATTTTTTTTGACTAC 363  
Qy 796 TTTGAGGCTGGAGCTACTCTTTTGACAGACAAGTCGATTATGGCTATTTCTTTCTGG 855  
Db 364 TTTGAGGCTGGAGCTACTCTTTTGACAGACAAGTCGATTATGGCTATTTCTCATCTGG 423  
Qy 856 AATGACAAATGGACAATGCAAGTTTGTTCGAAGATCTTATGCTCTTTACCGCTCAGATTTT 915  
Db 424 AATGACAAATGGACAAGGAGGATTCGTCCAAGATCTTATGCTCTTTACCGCTCAGACTTT 483  
Qy 916 TTTCCCGGTCTTGGATGGATGCTTTTCAAATCTACTTTGGACGAATATCTCCAAAGTGG 975



Db 484 NTCTCGTCTTGGATGGATGCTTTCAAATCAACTTGGTCCGAACATATCTCCAAAGTGG 543

Qy 976 CCAAGGCTTACTGGGACGACTGGCTAAGAC 1006

Db 544 CCGAAGGCTTACTGGGATGACTGGCTGAGGC 574

RESULT 14

LOCUS AJ806862 758 bp mRNA linear EST 08-DEC-2004

DEFINITION AJ806862 Antirrhinum majus whole plant Antirrhinum majus cDNA clone 018 6.05 c08, mRNA sequence.

ACCESSION AJ806862.1 GI:51122190

VERSION AJ806862.1

KEYWORDS EST.

ORGANISM Antirrhinum majus (snapdragon)

Antirrhinum majus

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamiales; Plantaginales; Antirrhineae; Antirrhinum.

REFERENCE 1 (bases 1 to 758)

AUTHORS Bey, M., Stueber, K., Fellenberg, K., Schwarz-Sommer, Z., Sommer, H., Siedler, H. and Zachgo, S.

TITLE Characterization of Antirrhinum Petal Development and Identification of Target Genes of the Class B MADS Box Gene DEFICIENS

JOURNAL Plant Cell 16 (12), 3197-3215 (2004)

PUBMED 15539471

COMMENT Contact: Schwarz-Sommer Z  
Molekulare Pflanzen-genetik  
MPI fuer Zuechtungs-forschung  
Carl-von-Linne Weg 10, D-50829, Germany.  
Location/Qualifiers  
1. 758  
/organism="Antirrhinum majus"  
/mol\_type="mRNA"  
/db\_xref="taxon:4151"  
/clone="018 6.05 c08"  
/tissue\_type="whole plant"  
/clone\_lib="Antirrhinum majus whole plant"

ORIGIN

Query Match 29.4%; Score 510.6; DB 1; Length 758;  
Best Local Similarity 82.0%; Pred. No. 2.8e-134;  
Matches 588; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

Qy 552 CATATCCAGGATGGATCACCTCGATGCTCAGGAAGCTTGTCTTGAGCTATGATCAGCT 611

Db 42 CGTATCACAGATGGATCAAAATCTGATGTTAAAGCATGGCGTTGAGTTATGATCAGCT 101

Qy 612 GAGCTATATGACGACTTGGATTTTGAACCTGTGCATCTGAAAGACCGAGGAGCTGAT 671

Db 102 AACCTACATGACGACTTGGACTATGACAGATCCGTACAGAAAGCGCTGGCGAGTTGAT 161

Qy 672 TGCATACTACAAATTTGACGCTATTACAGTGGGCAATTTGGATCGCTTTTACAGCA 731

Db 162 CGCATACTACAAAGATGACGCGCATTTACAAATGGGCACTTGATCAATTTTACAAACA 221

Qy 732 TAATTTAGCCGTGTATCATCTACAGATGATGGAATTTGCCCTGATTTTTTGA 791

Db 222 TAAATTTAGCCGATGATCATCTCGAAGATGATGGAATTTGCCCTGATTTCTTGA 281

Qy 792 CTTTTTTAGGCTGGAGCTACTCTCTTGACAGACAAAGTCGATTTATGGCTATTTCTTC 851

Db 282 GTATTTTGGCGGAGCGGCACTTCTTGACCGTGACAAATCCATCATGCAATTTTCATC 341

Qy 852 TTGGAATGACAAATGGACAAATGAGTTGTCCAAGATCTTATGCTCTTTACCGCTCAGA 911

Db 342 ATGGAATGATATGGAACACGCGAGTTTGTACATGATCTTATGCTGTACCGCTCAGA 401

Qy 912 TTTTTCCTCCGCTCTGGATGGATGCTTTCAAATCTACTTGGAGCAATATCTCCAAA 971

Db 402 TTTTTCCCTCGTCTTGGTGGATGCTTTTAAATCCACGTGGATGAATGTCTCTCTAA 461

Qy 972 GTGSCCAAAGGCTTACTTGGGACGACTGGCTAAGACTCAAAGAGAAATCACAGAGTCCGACA 1031

Db 462 ATGCCCCAAAGCTTACTTGGGATGACTGTTGCGACTCAAAGAAATCACAAAGGACGGCA 521

Qy 1032 ATTTATTCGCCCAAGCTTTTGCAGAACATATAATTTTGGTGAGCATGGTCTTAGTTGGG 1091

Db 522 ATTCAATTCGCCCAAGCTATGCAACATATAATTTTGGGAGCATGGATCCAGCATGGG 581

Qy 1092 GCAGTTTTTCAAGCAGTATCTTGAAGCAATTAATAAATGATCTCCAGGTTGATTTGAA 1151

Db 582 GCAATCTTCAACAGATATCTAGAACCTATCAAGATGAATGATATCCAGGTTGATTTGAA 641

Qy 1152 GTCAATGGACCTTAGTTTACCTTTTGGAGGACAAATTAAGTCTGAAACACCTTTGGTGACTTGGT 1211

Db 642 ATCAATGACCTCAGCTATCTGGAGGAGGACAGTATGTGAAGCACITTTGCTGACTTGTCT 701

Qy 1212 TAAAAAGGCTAAGCCCATCCATGGAGTGATGCTGTCTTGAAGCATTTAAACATAGA 1268

Db 702 TAAAAAGCGAGCGCTGTGCATGGGCTGCTGCTCTGAAAGGCAAAATATATAGA 758

RESULT 15

LOCUS BM780056

DEFINITION BM780056 852 bp mRNA linear EST 04-MAR-2002

ACCESSION EST590632 KV2 Medicago truncatula cDNA clone pkV2-53H18, mRNA

VERSION BM780056.1 GI:19110092

KEYWORDS EST.

SOURCE Medicago truncatula (barrel medic)

ORGANISM Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.

REFERENCE 1 (bases 1 to 852)

AUTHORS VandenBosch, K., Endre, G., Hur, J., Beremand, P., Town, C.D., Van Aken, S., Utterback, T., Cheung, F., Tsai, J. and Frazer, C.M.

TITLE ESTs from roots of Medicago truncatula 48 hr after inoculation with Sinorhizobium meliloti

JOURNAL Unpublished (2002)

COMMENT Contact: VandenBosch K  
Department of Plant Biology  
University of Minnesota  
220 Biosci Center, 1445 Gortner Ave, St. Paul, MN 55108, USA  
Tel: 612 624 2755  
Fax: 612 625 1738  
Email: kvandenb@cs.umn.edu  
TIGR sequence name: MTACT45TK More information is available at:  
www.medicago.org  
Seq primer: SKmod (CTA GAA CTA gtc gat CC).  
Location/Qualifiers  
1. 852  
/organism="Medicago truncatula"  
/mol\_type="mRNA"  
/cultivar="genotype A17"  
/db\_xref="taxon:3880"  
/clone="pkV2-53H18"  
/tissue\_type="Seedling roots"  
/dev\_stage="2 days post-inoculation with Sinorhizobium meliloti"  
/lab\_host="E. coli strain SOLR"  
/clone\_lib="KV2"  
/note="Vector: pBluescript SK -; Site 1: EcoRI; Site 2: XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA was directionally ligated into the Unizap XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-zap phage using Ex-assist helper phage and propagated in SOLR cells."

FEATURES

source

ORIGIN

Query Match	29.3%;	Score 508.2;	DB 3;	Length 852;
Best Local Similarity	75.3%;	Pred. No. 1.4e-133;		
Matches 633;	Conservative 0;	Mismatches 208;	Indels 0;	Gaps 0;
Qy	142	TTTTGCTGTGATTTCCGGTACCTCTCTCATCTTGGCTGCTGTGCGCTTCATCTACACACAG	201	
Db	2	TTTTGGGTGTATTTTCGGTGTCTCTCTTGTGTGCAGCCTTGGTGTTCACTCTACATCCAG	61	
Qy	202	ATCGCGCTTTTTCGGACACAGCTCAGAAATATGCAGATCGCCTTGTCTGCTGCAATTGAAGCA	261	
Db	62	ATGAGACTTTTCGCATCACAATCACAATATGCTGATCGCTCGCTGCTGCTATTGAAGCT	121	
Qy	262	GAATAATCATGTGTACAGCCAGACAGATGTCTTATTGACAGATTAGCTGTCGAGCAAGGA	321	
Db	122	GAATAATCACTGTACAGCTCAAAATCCGATCACTCATTTGATCAGATTAGCTTGCACAAGGA	181	
Qy	322	AGATATGTTGCTCTTTGAAGAACAAATGAAGCGTCAGGACACAGGAGTCCGCAATTAAGG	381	
Db	182	CGCATTTCTCGACTTACACAAGAAAGAACCGTCGAGAGCAAGAAATGTAGCCAAATCAAG	241	
Qy	382	GCTCTTGTTCAGGATCTTTGAAAGTAAGGGCATATAAAAGTTTGATCGGAAATGTACAGATG	441	
Db	242	TCTCTTGTTCAGATCTTTGAAGAAAAGATGTACGAAGCTGTATGATAAAGTCCAGGTT	301	
Qy	442	CCAGTGGCTGTGTAGTTGTTATTGGCTTTCGAATCGGGCTGATTTACCTGGAAAGAACTATT	501	
Db	302	CCAGTGGCAGCTGTTTGTTCATCATGTCATGTAATCGTGTGATTACCTGGAGAGGACTATC	361	
Qy	502	AAATCCCATCTTAAATACCAATATCTCTTGGGTCAAAATATCCTCTTTTCATATCCACG	561	
Db	362	AATTCTGTATTGAAATATCAAAAGGCCCAATTTCTTCAAGATTTCCTTTATTGTATCAAG	421	
Qy	562	GATGGATCACATCTCTGATGTTCAGGAAGCTTGCTTTGAGCTATGATCAGCTGACGTATATG	621	
Db	422	GATGGATCAAAATTCAGAGCTCAAAGAAAGGCTCTGAGCTACGATGAGCTATCTCATATG	481	
Qy	622	CAGCACTTGGATTTTGAACCTGTGTGCATATCTGAAAGACAGGGGAGCTGATTCATCTAC	681	
Db	482	CAGCACTTAGATTTTTCAGCCGGTTCAAACTGAAAGGCCAGGAGAGTTAATTGCTTACTAT	541	
Qy	682	AAATATGCACCTCATTAACAAGGGCATTTGGATCAGCTGCTTTTACAAGCATAAATTTAGC	741	
Db	542	AAGATTGCACTGTATTAACAAGTGGGCTCTTGGTCAACTGTTTCGACAAGCATAACTCAAC	601	
Qy	742	CGTGTATTACACTAGAGATGATATGAAATTTGCCCTTGATTTTTCGACTTTTTCGAG	801	
Db	602	AGAGTCATCATCTCGAAGATGACATGGAATAGCACCTGATTTCTTTGATTATTTCGA	661	
Qy	802	GCTGGAGCTACTCTTTTCGACAGAGACAAGTCGATATTGCGCTATTTCTTTTCGGAATGAC	861	
Db	662	GCTATGGCAACTCTCTTCGACAGGATTAAGTCCATTATGGCTGTTTCGTCATGGAATGAC	721	
Qy	862	AATGGACAAATGTCAGTTTGTCCAGATCCTTATGCTCTTTTACCGCTCAGATTTTTCCTCC	921	
Db	722	AATGGACAGAAGCAGTTTGTATCATGATCCATATGAACCTTATCGCTCGGACTTCTTCTCT	781	
Qy	922	GGTCTTGGATGGATGCTTTCAAAATCTACTTTGGGACGAATTATCTTCCAAAGTGGCCAAAG	981	
Db	782	GGACTAGGATGATGCTGGCTAGATCTACATGGGATGAACATATCACCAAAATGGCCCGAA	841	
Qy	982	G 982		
Db	842	G 842		

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2005, 06:18:00 ; Search time 8626 Seconds  
(without alignments)  
11446.450 Million cell updates/sec

Title: US-10-615-144-3

Perfect score: 1737  
Sequence: 1 gaattcgggcgccattga.....ataatcgggcgccgaattc 1737

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb.ba.\*  
2: gb.in.\*  
3: gb.env.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pr.\*  
9: gb.ro.\*  
10: gb.sts.\*  
11: gb.sy.\*  
12: gb.un.\*  
13: gb.vi.\*  
14: gb.htg.\*  
15: gb.pl.\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1737	100.0	1737	6	A95055 Sequence 3
2	1737	100.0	1737	6	BD226905 Plant Gnt
3	1737	100.0	1737	6	AR432619 Sequence
4	1708	98.3	1708	15	NTA249883 Nicotiana
5	1561.8	89.9	2031	15	NTY16832 Nicotiana t
6	1492.6	85.9	1836	15	NTA249882 Nicotiana
7	1297.6	74.7	1341	15	NBE295993 Nicotiana
8	1255	72.3	1641	15	STU249878 Solanum t
9	1255	72.3	1669	6	A95053 Sequence 1
10	1255	72.3	1669	6	BD226904 Plant Gnt
11	1255	72.3	1669	6	AR432618 Sequence
12	1239.2	71.3	1719	15	BT013875 Lycopersi
13	1216.6	70.0	1691	15	STU249879 Solanum t
14	1078	62.1	1499	15	STU249880 Solanum t
15	772	44.4	1625	15	BT000334 Arabidops
16	772	44.4	1766	15	AY099838 Arabidops
17	770.6	44.4	1820	15	ATH243198 Arabidops
18	767.2	44.2	1830	15	ATH249881 Arabidops

19	767.2	44.2	1854	6	A95057 Sequence 5
20	767.2	44.2	1854	6	BD226906 Plant Gnt
21	767.2	44.2	1854	6	AR432620 Sequence
22	688.2	38.5	1329	15	OS4457976 Oryza sat
23	688.2	38.5	1897	15	AK101526 Oryza sat
24	658.6	37.9	2180	6	AX840670 Sequence
25	658.6	37.9	9240	6	AX840669 Sequence
26	658	37.9	1319	6	AX840672 Sequence
27	656.6	37.8	1119	6	AX840688 Sequence
28	425.8	24.5	896	15	AK058826 Oryza sat
29	419.6	24.2	1416	15	AJ429143 Physcomit
30	417.2	24.0	422	15	AJ295994 Nicotiana
31	183.4	10.6	2695	9	NTA295994 Mus muscu
32	181.8	10.5	1894	9	MUSTRANX Mouse nuce
33	181.8	10.5	2651	9	MUSGLCNACT Mus muscu
34	181.8	10.5	203286	9	BC006629 Mouse DNA
35	181.6	10.5	2546	9	AL606829 Rat mRNA fo
36	181.6	10.5	2557	6	RATNAGT cDNA encodi
37	181.6	10.5	2595	9	E07296 Rat muscu
38	180.2	10.4	2988	9	BC074010 Rattus no
39	173.6	10.0	1344	9	BC031752 Mus muscu
40	172.8	9.9	2057	5	AF087456 Mesocrice
41	172	9.9	1344	9	BC088510 Xenopus t
42	165.8	9.5	1519	2	AF087457 Mesocrice
43	165.8	9.5	2894	5	AK115237 Ciona int
44	164.6	9.5	1344	9	X83975 X.laavis mr
45	164.2	9.5	2243	5	CGU65792 Cricetus
					BC041180 Xenopus 1

#### ALIGNMENTS

#### RESULT 1

A95055 1737 bp DNA linear PAT 26-JAN-2000  
Sequence 3 from Patent WO929879.  
A95055  
A95055.1 GI:6779208  
KEYWORDS  
SOURCE Nicotiana tabacum (common tobacco)  
ORGANISM  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
asterids; lamiales; Solanales; Solanaceae; Nicotiana.  
REFERENCE  
1 (bases 1 to 1737)  
AUTHORS  
TITLE  
VEGETABLE GNTI SEQUENCES AND THE USE THEREOF TO OBTAIN PLANTS WITH  
A REDUCED OR LACK OF N-ACETYLGLUCOSAMINYLTRANSFERASE I (GNTI)  
ACTIVITY  
JOURNAL  
Patent: WO 929879-A 1 17-JUN-1999;  
VON SCHAEWEN ANTJE (DE)  
FEATURES  
Location/Qualifiers  
1..1737  
/organism="Nicotiana tabacum"  
/mol\_type="unassigned DNA"  
/cultivar="Samsun NN"  
/db\_xref="taxon:4097"  
/cell\_type="BLATTZELLEN"  
/tissue\_type="MESOPHYLL"  
/clone\_lib="LAMBDA ZAP II (SCORI)"  
/dev\_stage="M: 'SINK' ORGAN"  
15..126  
5'UTR  
127..1467  
gene  
/gene="CGL"  
127..1467  
CDS  
/gene="CGL"  
/EC\_number="2.4.1.101"  
/function="INITIIERT KOMPLEXE N-GLYKANE AUF SEKRET.  
GLYKOPROTEINEN"  
/standard\_name="GNTI"  
/note="ERSTE GNTI-SEQUENZ AUS TABAK (UNPUBLIZIERT)"  
/codon\_start=1  
/evidence=experimental

		/product="BETA-1,2-N-ACETYLGLUCOSAMINYLTRANSFERASE I"	
		/protein_id="CAB70464.1"	
		/db_xref="GI:6779209"	
		/translation="MRGNKFCDDPRYLLILAAVAFIYQMLRFATSEYADRLAAALIE	
		AENHCTGOTRLIDQISLOQRIVALSEOMKRODOECQRLRALVQDLSEKIGKLIGN	
		VOMPVAAVVMACNRADYLEKTIKSLKYQISVASKYPLFISODGSHPDVVKIALSYD	
		QLTYMQLDPEPVHTERPEGLIAYKIRYKWAIDQLFYKHNFSRVILIEDDEIAP	
		DFFDFBAGATLLDRDKSIMAISSWNGNQMFQVDYALYRSDFPFGLGWLMSKSTW	
		DEUSPKWPKAYWDLRLKENHRGRQFIRPEVCRTYNFNGSHGSLGQFFKYLEPIKL	
		NDQVDKMSDLSYLLDEDYVHKFDLVKKAKPTIHGADAVLKAFNIDGDVRIQYRDQL	
		DFENIAKQFQIFEEWKDVPRAAYKGIWFRYQTSRRVFLVGHDSLQQLGIEDT"	
		<154..>213	
CDS		/gene="CGL"	
		/function="MEMBRANANKER (AMINO ACID 10-29) "	
		/standard_name="MEMBRANANKER EINES TYP II GOLGI-PROTEINS"	
		/note="DURCH VERGLEICH MIT TIERISCHEN GNTI-SEQUENZEN	
		IDENTIFIZIERT"	
		/codon_start=1	
		/product="HYDROPHOBER AMINO ACIDBEREICH IN GNTI"	
		/protein_id="CAB70465.1"	
		/db_xref="GI:6779210"	
		/translation="FRYLLILAAVAFIYQMLRF"	
		1468..1723	
3'UTR			
ORIGIN			
		Query Match 100.0%; Score 1737; DB 6; Length 1737;	
		Best Local Similarity 100.0%; Pred. No. 0;	
		Matches 1737; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1	GAATTCGGCGCGCCATTGACTCTCACTCACTCAACAGCAGCAAGTAATCCAGCGATGA	60
Db	1	GAATTCGGCGCGCCATTGACTCTCACTCACTCAACAGCAGCAAGTAATCCAGCGATGA	60
Qy	61	AACACTCATACTGAACACTGAGAGACTATTTCGCTTCTCTCTAAAGCCTTCAATCGGAATT	120
Db	61	AACACTCATACTGAACACTGAGAGACTATTTCGCTTCTCTCTAAAGCCTTCAATCGGAATT	120
Qy	121	CGCAGTAGAGAGGGAACAAGTTTGTGTGTAATTCGGGTACCTCTCATCTTGGCTGCT	180
Db	121	CGCAGTAGAGAGGGAACAAGTTTGTGTGTAATTCGGGTACCTCTCATCTTGGCTGCT	180
Qy	181	GTCCGCTTCATCTACACAGATCGGCTTTTTCGGACAGCTCAGATGAGATGAGATCGC	240
Db	181	GTCCGCTTCATCTACACAGATCGGCTTTTTCGGACAGCTCAGATGAGATGAGATCGC	240
Qy	241	CTTGCTGCTGCAATTGAAGCAGAAAATCAATGTGTACAGAGCAGATGAGATGAGATG	300
Db	241	CTTGCTGCTGCAATTGAAGCAGAAAATCAATGTGTACAGAGCAGATGAGATGAGATG	300
Qy	301	CAGATTAGCTGCGAGCAGGAAGAAATAGTGTCTCTTGAAGAACAAATGAAGCGTCAGGAC	360
Db	301	CAGATTAGCTGCGAGCAGGAAGAAATAGTGTCTCTTGAAGAACAAATGAAGCGTCAGGAC	360
Qy	361	CAGAGTGCCGACAAATTAAGGGCTCTTGTTCAGAGATCTTGAAGTAAGGGCATAAAG	420
Db	361	CAGAGTGCCGACAAATTAAGGGCTCTTGTTCAGAGATCTTGAAGTAAGGGCATAAAG	420
Qy	421	TTGATCGGAATGTACAGATGCCAGTGGCTGCTGTAGTTGTATTGGCTTGCATTCGGGCT	480
Db	421	TTGATCGGAATGTACAGATGCCAGTGGCTGCTGTAGTTGTATTGGCTTGCATTCGGGCT	480
Qy	481	GATTACCTGGAAAAGACTATTAAATCCATCTTAAATAACCAATATCTGTTCGCTCAAAA	540
Db	481	GATTACCTGGAAAAGACTATTAAATCCATCTTAAATAACCAATATCTGTTCGCTCAAAA	540
Qy	541	TATCCTCTTTTCATATCCAGGATGGATCAATCCTGATGTGAGGAGCTTGTCTTCAGC	600
Db	541	TATCCTCTTTTCATATCCAGGATGGATCAATCCTGATGTGAGGAGCTTGTCTTCAGC	600
Qy	601	TATGATCAGCTGATATATGAGCAGCTTGGATTTCGAACTGTGATCTGTAAGACCA	660
Db	601	TATGATCAGCTGATATATGAGCAGCTTGGATTTCGAACTGTGATCTGTAAGACCA	660

Qy	661	GGGGAGCTGATTCGCATACATACTACAAAATTGCAAGTCATTACAAAGTGGGCAATTGGATCAGCTG	720
Db	661	GGGGAGCTGATTCGCATACATACTACAAAATTGCAAGTCATTACAAAGTGGGCAATTGGATCAGCTG	720
Qy	721	TTTTCACAGCATAAATTTTAGCCGCTGTATCATACATAGAAAGATGATATGAAAATTTGCCCT	780
Db	721	TTTTCACAGCATAAATTTTAGCCGCTGTATCATACATAGAAAGATGATATGAAAATTTGCCCT	780
Qy	781	GAATTTTGTGACTTTTTTCAGGCTGGAGCTACTCTTCTTGACAGACACAAGTCGATATG	840
Db	781	GAATTTTGTGACTTTTTTCAGGCTGGAGCTACTCTTCTTGACAGACACAAGTCGATATG	840
Qy	841	GCTATTCTCTCTGGATGACATGCAATGCAAAATGCGAGTTTGTCCAAAGATCCTTATGCTCTT	900
Db	841	GCTATTCTCTCTGGATGACATGCAATGCAAAATGCGAGTTTGTCCAAAGATCCTTATGCTCTT	900
Qy	901	TACCGCTCAGATTTTTTTCGCCGCTTTCAGATGATGCTTTCAAAATCTACTTCGGACGAA	960
Db	901	TACCGCTCAGATTTTTTTCGCCGCTTTCAGATGATGCTTTCAAAATCTACTTCGGACGAA	960
Qy	961	TTATCTCCAAAGTGGCCAAAGGCTTACTCGGACGACTGGCTAAGACTCAAAAGAAATCAC	1020
Db	961	TTATCTCCAAAGTGGCCAAAGGCTTACTCGGACGACTGGCTAAGACTCAAAAGAAATCAC	1020
Qy	1021	AGAGGTGCAACATTTTATTCGCCCAGAAAGTTTGCAGAACATATAATTTTGGTGAGCATGGT	1080
Db	1021	AGAGGTGCAACATTTTATTCGCCCAGAAAGTTTGCAGAACATATAATTTTGGTGAGCATGGT	1080
Qy	1081	TCTAGTTTGGGGCAGTTTTTCAAGCAGTATCTTGAGCCAAATTAACCTAAATGATGTCAG	1140
Db	1081	TCTAGTTTGGGGCAGTTTTTCAAGCAGTATCTTGAGCCAAATTAACCTAAATGATGTCAG	1140
Qy	1141	GTGTGATTGGAAGTCAAATGGAACCTTAGTACTTTTCGAGAGCAAAATTCGTAAGAACATTT	1200
Db	1141	GTGTGATTGGAAGTCAAATGGAACCTTAGTACTTTTCGAGAGCAAAATTCGTAAGAACATTT	1200
Qy	1201	GGTGACTTGGTTAAAAGGCTAAGCCCATCCATCGAGCTGATGCTGCTTTGAAAGCATTTT	1260
Db	1201	GGTGACTTGGTTAAAAGGCTAAGCCCATCCATCGAGCTGATGCTGCTTTGAAAGCATTTT	1260
Qy	1261	AACATAGATGGTGTGATGCGTATTTCAGTACAGAGATCAACTAGACTTTTGAAAATATCGCA	1320
Db	1261	AACATAGATGGTGTGATGCGTATTTCAGTACAGAGATCAACTAGACTTTTGAAAATATCGCA	1320
Qy	1321	CGSCAAATTTGGCAATTTTGAAGAAATGGAAGGATGGTGTACCAAGTCGACGATATAAAGGA	1380
Db	1321	CGSCAAATTTGGCAATTTTGAAGAAATGGAAGGATGGTGTACCAAGTCGACGATATAAAGGA	1380
Qy	1381	ATAGTAGTTTTCGGGTACCAAGCTCCAGACGTGATTCCTTGTTCGCCATGATTCGCTTT	1440
Db	1381	ATAGTAGTTTTCGGGTACCAAGCTCCAGACGTGATTCCTTGTTCGCCATGATTCGCTTT	1440
Qy	1441	CAACAACTCGGAATTTGAAGATACCTTAAACAAAGATATGATTCGAGGAGCCCGGCAAAAT	1500
Db	1441	CAACAACTCGGAATTTGAAGATACCTTAAACAAAGATATGATTCGAGGAGCCCGGCAAAAT	1500
Qy	1501	TTTGACTTATTTGGGTAGGATGCAATCGAGCTGACACATAAACCATGATTTTACAGTTACAT	1560
Db	1501	TTTGACTTATTTGGGTAGGATGCAATCGAGCTGACACATAAACCATGATTTTACAGTTACAT	1560
Qy	1561	ACAAAGTTTATTTATCGGAGGAGCTCACTGTTCTAGTGTGTAAGGGATATCGGCTT	1620
Db	1561	ACAAAGTTTATTTATCGGAGGAGCTCACTGTTCTAGTGTGTAAGGGATATCGGCTT	1620
Qy	1621	CTTAGTATTGGATGAATCATCAACAAACCTATTATTTTAAAGTGTTCAGAACATAAAGAG	1680
Db	1621	CTTAGTATTGGATGAATCATCAACAAACCTATTATTTTAAAGTGTTCAGAACATAAAGAG	1680
Qy	1681	GAAATGAGCCCTGTAAAGACTATACATGGGACCATCATATCGCGCCCGCGAATTC	1737
Db	1681	GAAATGAGCCCTGTAAAGACTATACATGGGACCATCATATCGCGCCCGCGAATTC	1737

RESULT 2	BD226905	1737 bp	DNA	linear	PAT 17-JUL-2003
LOCUS	Plant GntI sequence and method of using the same for producing				
DEFINITION	plant with lowered or no N-acetylglucosaminyl transferase I (GntI) activity.				
ACCESSION	BD226905				
VERSION	BD226905.1	GI:33036675			
KEYWORDS	JP 2002509696-A/2.				
SOURCE	Nicotiana tabacum (common tobacco)				
ORGANISM	Nicotiana tabacum				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamids; Solanales; Solanaceae; Nicotiana.				
AUTHORS	1 (bases 1 to 1737)				
TITLE	Schaeven,A.V.				
JOURNAL	Plant GntI sequence and method of using the same for producing				
COMMENT	Patent: JP 2002509696-A 2 02-APR-2002.				
	ANTJUE VON SCHAEWEN				
	OS Nicotiana tabacum (tobacco)				
	PN JP 2002509696-A/2				
	PD 02-APR-2002				
	PF 09-DEC-1998 JP 2000524450				
	PR 09-DEC-1997 DE 197 54 622.6				
	PI ANTJUE VON SCHAEWEN				
	PC C12N15/09,A01H5/00,C07K16/40,C12N5/10,C12N9/10//C12P21/08, PC C12N15/00,				
	PC C12N5/00				
	CC Strandedness: Double;				
	CC Topology: Linear;				
	CC /function= 'Asn codon in this context is a potential CC				
	glycosylation site'				
	CC /product= 'N-glycosylation consensus sequence' CC				
	/phenotype= 'N-glycans modulate protein properties' CC				
	/standard name= 'N-glycosylation site'				
	CC /label= 'pot-CHO				
	CC /note= 'GntI sequences from animals do not contain this CC				
	feature'				
	CC /codon_start= 127				
	CC /functions= 'initiates complex N-glycans on secretory CC				
	glycoproteins'				
	CC /EC_number= 2.4.1.101				
	CC /product= 'beta-1,2-N-acetylglucosaminyltransferase I' CC				
	/evidence= EXPERIMENTAL				
	CC /gene= 'cgl'				
	CC /standard_name= 'gntI'				
	CC /label= ORF				
	CC /note= 'first gntI sequence from tobacco (unpublished)' CC				
	/function= 'membrane anchor (amino acids 10-29)' CC /product=				
	'hydrophobic amino acid stretch in GntI' CC /standard name=				
	'membrane anchor of a type II golgi protein' CC /function= 'use				
	for cloning the cDNA library in Lambda ZAPII' CC /product=				
	'EcoRI/NotI-cDNA adapter'				
	CC /number= 1				
	CC /product= 'EcoRI/NotI-cDNA adapter'				
	CC /number= 2				
	PH Key	Location/Qualifiers			
	FT misc_feature	733..741			
	FT CDS	127..1467			
	FT 5'UTR	15..126			
	FT 3'UTR	1468..1723			
	FT CDS	154..213			
	FT misc_feature	1..14			
	FT misc_feature	1724..1737.			
	Location/Qualifiers				
	1..1737				
	/organism= 'Nicotiana tabacum'				
	/mol_type= 'genomic DNA'				
	/db_xref= 'taxon:4097'				
FEATURES					
source					
ORIGIN					
Query Match	100.0%;	Score 1737;	DB 6;	Length 1737;	

Dn	1021	AGAGGTGCAAAATTTATTTCGCCAGAAAGTTTGAGAACATATATATTTTGGTGACATGGT	1080
Qy	1081	TCTAGTTTGGGCGAGTTTTCAGCAGATATCTTGAGCCAAATTAACCTAAATGATGTCAG	1140
Dn	1081	TCTAGTTTGGGCGAGTTTTCAGCAGATATCTTGAGCCAAATTAACCTAAATGATGTCAG	1140
Qy	1141	GTTGATTTGGAAGTCAATGGACCTTAGTTACCTTTGGAGGACAAATTAAGTGAACACCTTT	1200
Dn	1141	GTTGATTTGGAAGTCAATGGACCTTAGTTACCTTTGGAGGACAAATTAAGTGAACACCTTT	1200
Qy	1201	GGTGACTTGGTTAAAAAGGCTTAAGCCCATCCATGGAGCTGATGCTGCTTTGAAAGCAATTT	1260
Dn	1201	GGTGACTTGGTTAAAAAGGCTTAAGCCCATCCATGGAGCTGATGCTGCTTTGAAAGCAATTT	1260
Qy	1261	AACATAGATGGTATGTCGTAATTCAGTACAGAGATCAACTAGACTTTGAAAAATATCGCA	1320
Dn	1261	AACATAGATGGTATGTCGTAATTCAGTACAGAGATCAACTAGACTTTGAAAAATATCGCA	1320
Qy	1321	CGGCAATTTGGCAATTTTGAAGATGGAAGATGGTGTACCCAGTGCAGCATATAAAGGA	1380
Dn	1321	CGGCAATTTGGCAATTTTGAAGATGGAAGATGGTGTACCCAGTGCAGCATATAAAGGA	1380
Qy	1381	ATAGTAGTTTCCGGTACCAACCTCCAGAGCTGATTTCTTGTGGCCATGATTCGCTTT	1440
Dn	1381	ATAGTAGTTTCCGGTACCAACCTCCAGAGCTGATTTCTTGTGGCCATGATTCGCTTT	1440
Qy	1441	CAACAACCTCGGAATGGAAGATCTTAAACAAAGATATGATTCAGAGGCGCGGGCAAAATTT	1500
Dn	1441	CAACAACCTCGGAATGGAAGATCTTAAACAAAGATATGATTCAGAGGCGCGGGCAAAATTT	1500
Qy	1501	TTTGACTTATTTGGGTAGGATGCATCGAGCTGACACTAAACCATGATTTTACCAGTTACAT	1560
Dn	1501	TTTGACTTATTTGGGTAGGATGCATCGAGCTGACACTAAACCATGATTTTACCAGTTACAT	1560
Qy	1561	ACAAAGTTTAAATGTTATACGGAGAGCTCACTGTTCTAGTGTGTAAGGGATATCGGCTTT	1620
Dn	1561	ACAAAGTTTAAATGTTATACGGAGAGCTCACTGTTCTAGTGTGTAAGGGATATCGGCTTT	1620
Qy	1621	CTTAGTATGGAATGATCAACACCAACCTATTATTTAAAGTGTTCAGAACATAAAGAG	1680
Dn	1621	CTTAGTATGGAATGATCAACACCAACCTATTATTTAAAGTGTTCAGAACATAAAGAG	1680
Qy	1681	GAAATGTAGCCCTTAAAGACTATACATGGGACCATCATATTCGGCGCGGAATTC	1737
Dn	1681	GAAATGTAGCCCTTAAAGACTATACATGGGACCATCATATTCGGCGCGGAATTC	1737

RESULT 3  
AR432619 LOCUS AR432619 1737 bp DNA linear PAT 18-DEC-2003  
DEFINITION Sequence 3 from patent US 6653459.  
ACCESSION AR432619  
VERSION AR432619.1 GI:40195152  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1737)  
AUTHORS Von Schaeuwen,A.  
TITLE Plant Gnt1 sequences and the use thereof for the production of plants having reduced or lacking N-acetyl glucosaminyl transferase I(Gnt1) activity  
JOURNAL Patent: US 6653459-A 3 25-NOV-2003;  
FEATURES Location/Qualifiers  
source 1..1737  
/organism="unknown"  
/mol\_type="genomic DNA"

= parent of instant app.

Query Match 100.0%; Score 1737; DB 6; Length 1737;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1737; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GAATTTCGGCGCCGCAATTTGACTTCCTAACTGAACAGGCAAGATAAATCCACGCGATGA	60
Dn	1	GAATTTCGGCGCCGCAATTTGACTTCCTAACTGAACAGGCAAGATAAATCCACGCGATGA	60
Qy	61	AACACTCATAACTGAACACACTGAGAGACTATTCCTTAAAGCCTTCAATCGAATTT	120
Dn	61	AACACTCATAACTGAACACACTGAGAGACTATTCCTTAAAGCCTTCAATCGAATTT	120
Qy	121	CGCAGATGAGAGGGAACAAGTTTTCCTGCTGATTTCCGGTACCTCTCATCTTGGCTGCT	180
Dn	121	CGCAGATGAGAGGGAACAAGTTTTCCTGCTGATTTCCGGTACCTCTCATCTTGGCTGCT	180
Qy	181	GTGCGCTTCATCTACACAGATGCGGCTTTTTCGACACAGTCAGATATGAGATCGC	240
Dn	181	GTGCGCTTCATCTACACAGATGCGGCTTTTTCGACACAGTCAGATATGAGATCGC	240
Qy	241	CTTGCTGCTGCAATTTGAAGCAGAAAAATCAATTTGTAAGCCAGACAGATTGCTTATTGAC	300
Dn	241	CTTGCTGCTGCAATTTGAAGCAGAAAAATCAATTTGTAAGCCAGACAGATTGCTTATTGAC	300
Qy	301	CAGATTAGCTGAGCAAGAAATAGTTGCTCTTGAAGAAACAATGAAGCGTCAGGAC	360
Dn	301	CAGATTAGCTGAGCAAGAAATAGTTGCTCTTGAAGAAACAATGAAGCGTCAGGAC	360
Qy	361	CAGGAGTGGCGCAATTTAAAGGCTCTTGTTCAGGATCTTGAAGTAAGGCGCATAAAAAG	420
Dn	361	CAGGAGTGGCGCAATTTAAAGGCTCTTGTTCAGGATCTTGAAGTAAGGCGCATAAAAAG	420
Qy	421	TTGATCGGAAATGTACAGATGCGGCTGCTGTAGTTGTTTATGGCTTCAATTCGGGCT	480
Dn	421	TTGATCGGAAATGTACAGATGCGGCTGCTGTAGTTGTTTATGGCTTCAATTCGGGCT	480
Qy	481	GATTACCTGGAAGAAGCTATTAAATCCATCTTAAATAACCAATATCTGTGGCTCAAAA	540
Dn	481	GATTACCTGGAAGAAGCTATTAAATCCATCTTAAATAACCAATATCTGTGGCTCAAAA	540
Qy	541	TATCTCTTTTTCATATCCAGGATGATCACATCTGATGTCAGGAGCTTGTCTTCAGC	600
Dn	541	TATCTCTTTTTCATATCCAGGATGATCACATCTGATGTCAGGAGCTTGTCTTCAGC	600
Qy	601	TATGATCAGCTGACGTATATGACGACCTTGGATTTTGAACCTGTGCATATCTGAAAGACCA	660
Dn	601	TATGATCAGCTGACGTATATGACGACCTTGGATTTTGAACCTGTGCATATCTGAAAGACCA	660
Qy	661	GGGAGCTGATTCGATACATACAAAAATTCAGCTCAATTAAGTGGGCAATTCGATCGCTG	720
Dn	661	GGGAGCTGATTCGATACATACAAAAATTCAGCTCAATTAAGTGGGCAATTCGATCGCTG	720
Qy	721	TTTTCAAGCATAAATTTTAGCCCTGTTATCATACTAGAGATGATATGGAATTTGCCCTT	780
Dn	721	TTTTCAAGCATAAATTTTAGCCCTGTTATCATACTAGAGATGATATGGAATTTGCCCTT	780
Qy	781	GATTTTGTGACCTTTTTCAGGCTGGAGCTACTCTTTCGACAGAGACAAGTCGATTATG	840
Dn	781	GATTTTGTGACCTTTTTCAGGCTGGAGCTACTCTTTCGACAGAGACAAGTCGATTATG	840
Qy	841	GCTATTTCTTCTTGGAAATGCAATGCAATGCAAGTTTTCGCAAGATTCCTTATGCTCTT	900
Dn	841	GCTATTTCTTCTTGGAAATGCAATGCAATGCAAGTTTTCGCAAGATTCCTTATGCTCTT	900
Qy	901	TACCGCTCAGATTTTTCCTCGGCTTGGATGATGCTTTCAAAATCTACTTGGAGCGAA	960
Dn	901	TACCGCTCAGATTTTTCCTCGGCTTGGATGATGCTTTCAAAATCTACTTGGAGCGAA	960
Qy	961	TTATCTCCAAAGTGGCCAAAGGCTTACTGGGACGCTGGCTTAAGACTCAAAGGAATCAC	1020
Dn	961	TTATCTCCAAAGTGGCCAAAGGCTTACTGGGACGCTGGCTTAAGACTCAAAGGAATCAC	1020
Qy	1021	AGAGGTGCAAAATTTATTCGCCAGAGGTTTTCGAGAACATATATTTTGGTGAGCATGGT	1080
Dn	1021	AGAGGTGCAAAATTTATTCGCCAGAGGTTTTCGAGAACATATATTTTGGTGAGCATGGT	1080
Qy	1081	TCTAGTTTGGGCGAGTTTTCAGCAGATATCTTGAGCCAAATTAACCTAAATGATGTCAG	1140



615 GTATATGCGACCTTGGATTTTGAACCTGTGTCATCTGNAAGACCGAGGAGCTGATTGC 674  
Db GTATATGCGACCTTGGATTTTGAACCTGTGTCATCTGNAAGACCGAGGAGCTGATTGC 660  
675 ATACTACAAAATTCGACGTCATTAACAAGTGGCGATTCGATCAGCTGTTTTCACAAGCATAA 734  
Db ATACTACAAAATTCGACGTCATTAACAAGTGGCGATTCGATCAGCTGTTTTCACAAGCATAA 720  
735 TTTTAGCGGTGTTATCATATACATAGAAATGATATGAAATTCGCCCTGATTTTGTGACTT 794  
Db TTTTAGCGGTGTTATCATATACATAGAAATGATATGAAATTCGCCCTGATTTTGTGACTT 780  
795 TTTTAGCGGTGACTACTCTCTTGACAGAGACAAGTCGATATGCTATTTCTTCTTCTTG 854  
Db TTTTAGCGGTGACTACTCTCTTGACAGAGACAAGTCGATATGCTATTTCTTCTTCTTCTTG 840  
855 GAATGACAATGCAAAATGCAATTTGTGCAAGATCTTATGCTCTTTACCGCTCAGATTT 914  
Db GAATGACAATGCAAAATGCAATTTGTGCAAGATCTTATGCTCTTTACCGCTCAGATTT 900  
915 TTTTCCCGCTTGGATGATGCTTTTCAAAATCTACTTGGACGAATATCTCCAAAGTG 974  
Db TTTTCCCGCTTGGATGATGCTTTTCAAAATCTACTTGGACGAATATCTCCAAAGTG 960  
975 GCCAAGGCTTACTGGGACGACTGGCTAAGACTCAAGAGACTCAAGAGATCAGAGGTGCAAAAT 1034  
Db GCCAAGGCTTACTGGGACGACTGGCTAAGACTCAAGAGACTCAAGAGATCAGAGGTGCAAAAT 1020  
1035 TATTGCGCCAGAAAGTTTGCAGAACATATATTTTGGTGAGCATGTTCTAGTTTGGGGCA 1094  
Db TATTGCGCCAGAAAGTTTGCAGAACATATATTTTGGTGAGCATGTTCTAGTTTGGGGCA 1080  
1095 GTTTTTCAGCAGATCTCTGAGCCAAATTAATAAATCATGTCCAGGTGATTTGGAAGTC 1154  
Db GTTTTTCAGCAGATCTCTGAGCCAAATTAATAAATCATGTCCAGGTGATTTGGAAGTC 1140  
1155 AATGACCTTAGTTACCTTTTGGAGGACAAATGATGTAACACACACACACACACACACACAC 1214  
Db AATGACCTTAGTTACCTTTTGGAGGACAAATGATGTAACACACACACACACACACACACAC 1200  
1215 AAAGGCTAAGCCCATCGATGAGCTGCTGCTTGAAGCATTTAAACATAGATGTTGA 1274  
Db AAAGGCTAAGCCCATCGATGAGCTGCTGCTTGAAGCATTTAAACATAGATGTTGA 1260  
1275 TGTGCGTATTTCAGTACAGAGATCAACTAGACTTTGAAATAATCGCAGCAATTTGGCAT 1334  
Db TGTGCGTATTTCAGTACAGAGATCAACTAGACTTTGAAATAATCGCAGCAATTTGGCAT 1320  
1335 TTTTGAAGATGGAAGGATGTTACCCAGCTGAGCATATAAGGAATAGTAGTTTCCG 1394  
Db TTTTGAAGATGGAAGGATGTTACCCAGCTGAGCATATAAGGAATAGTAGTTTCCG 1380  
1395 GTACCAACGCTCCAGACGCTGATTTCTTTTGGCCATGATTCGCTTCAACAACTCGGAAT 1454  
Db GTACCAACGCTCCAGACGCTGATTTCTTTTGGCCATGATTCGCTTCAACAACTCGGAAT 1440  
1455 TGAAGATACCTTAAACAAGATATGATTCAGGAGCCCGGCAAAATTTTGTGACTTATTGGG 1514  
Db TGAAGATACCTTAAACAAGATATGATTCAGGAGCCCGGCAAAATTTTGTGACTTATTGGG 1500  
1515 TAGATGATCGAGCTGACATCAACCATGATTTTACAGTTATACATACAGCTTTTAAATG 1574  
Db TAGATGATCGAGCTGACATCAACCATGATTTTACAGTTATACATACAGCTTTTAAATG 1560  
1575 TTATACGAGGAGCTCACTGTTCTAGTGTTCGAAGGATATCGGCTCTTAGTATTGGATG 1634  
Db TTATACGAGGAGCTCACTGTTCTAGTGTTCGAAGGATATCGGCTCTTAGTATTGGATG 1620  
1635 AATCATCAACAACTATTTTAAAGTGTTCAGAAATAAGAGGAAATGTAGCCCTG 1694  
Db AATCATCAACAACTATTTTAAAGTGTTCAGAAATAAGAGGAAATGTAGCCCTG 1680

QY 1695 TAAAGACTATACATGGGACCATCATAT 1722  
Db 1681 TAAAGACTATACATGGGACCATCATAT 1708  
RESULT 5  
NTY16832  
LOCUS  
DEFINITION  
NTY16832 2031 bp mRNA linear PLN 18-APR-2005  
Nicotiana tabacum mRNA for  
beta-1,2-N-acetylglucosaminyltransferase, complete CDS.  
Y16832  
ACCESSION  
VERSION  
Y16832.1 GI:5764048  
KEYWORDS  
beta-1,2-N-acetylglucosaminyltransferase.  
SOURCE  
Nicotiana tabacum (common tobacco)  
ORGANISM  
Nicotiana tabacum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
asterids; lamids; Solanales; Solanaceae; Nicotiana.  
REFERENCE  
1 Strasser, R., Mucha, J., Schwihla, H., Altmann, F., Glossl, J. and  
Steinkellner, H.  
Molecular cloning and characterization of cDNA coding for beta1,  
2N-acetylglucosaminyltransferase I (GlcNAc-TI) from Nicotiana  
tabacum  
Glycobiology 9 (8), 779-785 (1999)  
10408843  
REFERENCE  
2 (bases 1 to 2031)  
Steinkellner, H.  
Direct Submission  
TITLE  
Submitted (06-MAR-1998) H. Steinkellner, Zentrum fuer Angewandte  
Genetik, BOKU-Wien, Muthgasse 18, 1190 Wien, AUSTRIA  
JOURNAL  
FEATURES  
source  
1. .2031  
/organism="Nicotiana tabacum"  
/mol\_type="mRNA"  
/cultivar="Xanthi"  
/db\_xref="taxon:4097"  
1. .76  
77. .1417  
/EC\_number="2.4.1.101"  
/codon\_start=1  
/product="alpha-1,3-mannosyl-glycoprotein  
beta-1,2-N-acetylglucosaminyltransferase"  
/protein\_id="CA553347.1"  
/db\_xref="GI:5764048"  
/db\_xref="GOA:O9ST97"  
/db\_xref="InterPro:IPR001173"  
/db\_xref="InterPro:IPR004139"  
/db\_xref="UniProt/TREMBL:Q9ST97"  
/translation="MRGYKFCDFRYLLILAAVFIYIQMLFPATOSEYADRLAAALE  
AENHCTSQRLLIDQISQOGRIVALEQMKQDQECRQLRALVQDLESKIKKLIQGN  
VQMPVAVVVMACNRADYLEKTIKSIKYQISVAPKPLFISQDGHDPVRLKLIQSYD  
QUTMQLDFEPVHTERPGEIAYYKIAHYKWDQLFYKHNFSRVILIEDDMEIAP  
DFFDFEAGATLLDRKSIATSSWNDQMOFQVQYALYRSDFFPGLGWLMSKSTW  
DELSPKWPKAYWDDWLRLKENHRGROFIRPEVCSYNGEHSGLGQFPKQYLEPIKL  
NDVQDWKSNLSDYLLEDNYVHFGLVKAPDHGDAVLKAPNIDGVDRIQYRDOL  
FVEDIARQGFIEEMKDGVPRAAYKGIIVFVRYQTSRRVFLVGVGDSLQQUGNEDI"  
1418. .2031  
3' UTR  
ORIGIN  
Query Match 89.9%; Score 1561.8; DB 15; Length 2031;  
Best Local Similarity 96.8%; Pred. No. 0;  
Matches 1627; Conservative 0; Mismatches 47; Indels 7; Gaps 3;  
QY 51 CCACGAGTGAACACTCATAACTGAACACTGAGAGACTATTCGCTTTCTCTAAAGCCTT 110  
Db 1 CCACGAGTGAACACTCATAACTGAACACTGATAGACTATTCGCTTTCTCTAAAGCCTT 60  
QY 111 CAATCGAATTCGACGATGAGGGAAACAAGTTTTCGTGTGATTTCCGGTACCTTCAT 170  
Db 61 CAATCGAATTCGACGATGAGGGGTACAAGTTTTCGTGTGATTTCCGGTACCTTCAT 120  
QY 171 CTTGGCTGCTGCTGCCTTCATCTACACACAGATCGCGCTTTTTCGCGACACAGTCAAGAATA 230



Db 121 |||||CTTGGCTGCTGCGCTTCATCTACATACAGATCGGCTTTTTCGGACACAGTCAGAATA 180  
Qy 231 TGCAGATCGCTTCTGCTGCAATTTGAAGCGAAAAATCATTTGTCAAGCCAGACCAAGATT 290  
Db 181 TGCAGATCGCTTCTGCTGCAATTTGAAGCGAAAAATCATTTGTCAAGCCAGACCAAGATT 240  
Qy 291 GCTTATTGACAGATAGCTGCGCAAGGAAAGTAATAGTTCTCTTGAAGAAACAATGAA 350  
Db 241 GCTTATTGACAGATTTAGCCAGCGAAGAAAGTAATAGTTCTCTTGAAGAAACAATGAA 300  
Qy 351 GCGTCAGGACCAAGGATGCGCAAAATTAAGGGCTCTTGTTCAGGATCTTGAAGTAAGGG 410  
Db 301 GCGTCAGGACCAAGGATGCGCAAAATTAAGGGCTCTTGTTCAGGATCTTGAAGTAAGGG 360  
Qy 411 CATAAAAAGTTGATCGGAATGTACAGATGCCAGTGGCTGCTGTAGTTGTATGGCTTG 470  
Db 361 CATAAAAAGTTGATCGGAATGTACAGATGCCAGTGGCTGCTGTAGTTGTATGGCTTG 420  
Qy 471 CAATCGGCTGATTTACCTGGAAGAGACTATTAAATCCATCTTAAATACCAAAATATCTGT 530  
Db 421 CAATCGGCTGATTTACCTGGAAGAGACTATTAAATCCATCTTAAATACCAAAATATCTGT 480  
Qy 531 TCGCTCAAAATATCTCTTTTCAATCCAGGATGGATCACATCCTGTATGTCAGGAAGCT 590  
Db 481 TCGGCCAAATATCTCTTTTCAATCCAGGATGGATCACATCCTGTATGTCAGGAAGCT 540  
Qy 591 TCGTTTGAAGTATGATCAGCTGACGTATATGCGACCTTGGATTTTGAACCTGTGCATAC 650  
Db 541 TCGTTTGAAGTATGATCAGCTGACGTATATGCGACCTTGGATTTTGAACCTGTGCATAC 600  
Qy 651 TGAAGACACGGGAGCTGATTTGCACTACACAAATTCAGCTCATTAACAAGTGGGCATT 710  
Db 601 TGAAGACACGGGAGCTGATTTGCACTACACAAATTCAGCTCATTAACAAGTGGGCATT 660  
Qy 711 GGATCAGCTGTTTTACAAGCAATAATTTAGCGCTGTATCATCTAGAGATGATATGGA 770  
Db 661 GGATCAGCTGTTTTACAAGCAATAATTTAGCGCTGTATCATCTAGAGATGATATGGA 720  
Qy 771 AATGCCCTGATTTTTTGACTTTTTTGAGGCTGGAGCTACTCTTCTTGACAGAGACAA 830  
Db 721 AATGCCCTGATTTTTTGACTTTTTTGAGGCTGGAGCTACTCTTCTTGACAGAGACAA 780  
Qy 831 GTCGATTATGCTATTTCTTTGCAATGACAAATGCAAAATGCAAGTGTGTCAGAGATCC 890  
Db 781 GTCGATTATGCTATTTCTTTGCAATGACAAATGCAAAATGCAAGTGTGTCAGAGATCC 840  
Qy 891 TTATGCTCTTTACCGCTCAGATTTTTTCCCGCTCTTGGATGGATGCTTTCAAAATCTAC 950  
Db 841 TTATGCTCTTTACCGCTCAGATTTTTTCCCGCTCTTGGATGGATGCTTTCAAAATCTAC 900  
Qy 951 TTGGGACGAATTTATCTCCAAAGTGCCCAAGGCTTACTGGGACGACTGGCTAAGACTCAA 1010  
Db 901 TTGGGACGAATTTATCTCCAAAGTGCCCAAGGCTTACTGGGACGACTGGCTAAGACTCAA 960  
Qy 1011 AGAGAAATCACAGAGTCGCAATTTATTCGCCAGAGTTTTCAGAACATATAAATTTGG 1070  
Db 961 AGAGAAATCACAGAGTCGCAATTTATTCGCCAGAGTTTTCAGAACATATAAATTTGG 1020  
Qy 1071 TCAGCATGCTTCTAGTTTGGGCGAGTTTTTCAAGCAGATATCTTGAGCCAAATTAACCTAA 1130  
Db 1021 TCAGCATGCTTCTAGTTTGGGCGAGTTTTTCAAGCAGATATCTTGAGCCAAATTAACCTAA 1080  
Qy 1131 TGATGTCAGAGTTGATTTGGAAGTCAATGGACCTTAGTTACCTTTTGGAGGACAAATTAAGT 1190  
Db 1081 TGATGTCAGAGTTGATTTGGAAGTCAATGGACCTTAGTTACCTTTTGGAGGACAAATTAAGT 1140  
Qy 1191 GAAACACTTTTGTGATCTGGTTAAAGGCTAAGCCATCCATGGAGCTGATGCTGCTTT 1250  
Db 1141 GAAACACTTTTGTGATCTGGTTAAAGGCTAAGCCATCCATGGAGCTGATGCTGCTTT 1200  
Qy 1251 GAAAGCATTTTAAATAGATGCTGCTGATTTAGTACAGAGATCAACTAGACTTTCA 1310

Db 1201 GAAAGCATTTTAAATAGATGGTGAATGTCGTATTTCAAGTACAGAGATCAACTAGACTTTGA 1260  
Qy 1311 AAATATCGCACGGCAATTTTGGCATTTTGAAGAAATGAAGAGTGGTGTACACGTCAGC 1370  
Db 1261 AGATATCGCACGGCAATTTTGGCATTTTGAAGAAATGAAGAGTGGTGTACACGTCAGC 1320  
Qy 1371 ATATAAAGGAATAGTAGTTTTCGGGTACCAAAAGTCCAGACGTTATTCCTTGTGGCCA 1430  
Db 1321 ATATAAAGGAATAGTAGTTTTCGGGTACCAAAAGTCCAGACGTTATTCCTTGTGGCCC 1380  
Qy 1431 TGATTCGCTTCAACCACTCGGAATTTGAAGATCTTTAAACAAGATATGATTGTCAGAGAGCCC 1490  
Db 1381 TGATTCGCTTCAACCACTCGGAATTTGAAGATCTTTAAACAAGATATGATTGTCAGAGAGCCC 1437  
Qy 1491 GGGCAAAATTTTTCGCTTATTTGGTAGGATCGATCGAGCTGACACTAAACCATGATTTTA 1550  
Db 1438 GGACAAAGATTTAGACTTTATTTGGTAGGATCGATCGAGCTGACACTAAACCATGATTTTA 1497  
Qy 1551 CCAGTTACATACAAAGTTTAA--TGTTATACGGAGGAGCTCAGTCTTAGTGTGAAGG 1609  
Db 1498 CCAGTTACATACAAAGTTTAAATTTGTTATATGAGGAGCTCAGTCTTAGTGTGAAGG 1557  
Qy 1610 GATATCGCTTCTTAGTATTTGGATGAATCATCAACAACTATTATT---TTAAGTGT 1666  
Db 1558 GATATCGCTTCTTAGTATTTGGATGAATCATCAACAACTATTATT---TTAAGTGT 1617  
Qy 1667 CAGAACATAAAGAGAAATGTAGCCCTGTAAAGACTATACATGGGACCATCATATCCGC 1726  
Db 1618 CCGAACATAAAGAGAAATGTAGCCCTGTAAAGACTATACATGGGACCATCATATCCGC 1677  
Qy 1727 G 1727  
Db 1678 G 1678

RESULT 6  
NTA249882  
LOCUS Nicotiana tabacum mRNA for N-acetylglucosaminyltransferase I (GntI gene); clone A4. 1836 bp mRNA linear PLN 15-APR-2005

DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

REFERENCE  
AUTHORS  
TITLE

JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

FEATURES  
source  
1. 1836  
/organism="Nicotiana tabacum"  
/mol\_type="mRNA"  
/cultivar="Samsun NN"  
/db\_xref="taxon:4097"  
/clone="A4"  
/tissue\_type="leaf mesophyll"  
/note="sink"  
1. 1836  
/gene="GntI"  
27. 1367

gene  
CDS

```

/gene="GntI"
/EC_number="2.4.1.101"
/function="initiates complex-type glycans in the Golgi"
/codon_start=1
/product="N-acetylglucosaminyltransferase I"
/protein_id="CAC80701.1"
/db_xref="GI:18076146"
/db_xref="COA:Q8VXD4"
/db_xref="InterPro:IPR001173"
/db_xref="InterPro:IPR004139"
/db_xref="UniProt/TREMBL:Q8VXD4"
/translation="MRGKFCDFPRYLILAAVAFIYIOMLFATOSEYADRLAAIE
AENHCTOTRLIDIOISQOGRIVALEEQKRODOECROLRALVODLESQIKKLIGN
VQMPVAVVMACNRADYLEKTKSILKYQISVAPKYPLEISQDSHPDVVKLALSD
QLTMHLDPEFHPERBELIAYIKARHKWALQDLFYKHNPSRVILLEDMETAP
DFDFEAGATLLDRDKS IMALISWNDNGQMFQVDPYALYRSDFFPGLGMLSKSTW
DELSPKWPYDWDLRLKENHGRQFIRPEVCRSYNFGEHSGSLGQFPFKYLEPKL
NDVODWKSMDLSYLLDENVKHPGDLVKKAKPIHGADAVLKA FNIDGDVRIQVRDL
DFEDIAHQFGIFIEBQDGPRAAYKGIIVFRYQTSRRVFLVGPDSLQLQGNEDT"

```

ORIGIN

```

Query Match      85.9%; Score 1492.6; DB 15; Length 1836;
Best Local Similarity 96.7%; Pred. No. 0;
Matches 1579; Conservative 0; Mismatches 44; Indels 10; Gaps 5;

Qy 101 CTAAGCCTTCAATCGAATTCGACGATGAGAGGGAACAAGTTTGTCTGTGATTCCGGT 160
Db 1 CTAAGCCTTCAATCGAATTCGACGATGAGAGGGAACAAGTTTGTCTGTGATTCCGGT 60

Qy 161 ACCTCCCTCATCTTGGCTGCTGCGCTTCATCTACACAGATCGCGCTTTTTCGACAC 220
Db 61 ACCTCCCTCATCTTGGCTGCTGCGCTTCATCTACACAGATCGCGCTTTTTCGACAC 120

Qy 221 AGTCAGAAATGACAGATGCGCTTCTGCTGCAATTTGAAGCAGAGAAATCATTTGACAAAGC 280
Db 121 AGTCAGAAATGACAGATGCGCTTCTGCTGCAATTTGAAGCAGAGAAATCATTTGACAAAGC 180

Qy 281 AGACCAAGTTCCTTATTTGACCAAGATTTAGCTGCGACGAGGAGGAGATGTTGCTCTTGAAG 340
Db 181 AGACCAAGTTCCTTATTTGACCAAGATTTAGCTGCGACGAGGAGGAGGAGATGTTGCTCTTGAAG 240

Qy 341 AACAAATGAAGCGTCAGGACGAGGTCGCCACAATTAAGGCTCTTTGTCAGATCTTG 400
Db 241 AACAAATGAAGCGTCAGGACGAGGTCGCCACAAGTTAAGGCTCTTTGTCAGATCTTG 300

Qy 401 AAAGTAAGGGCATAAAAAAGTTGATCGGAAATGTACAGATGCCAGTGGCTGCTGTAGTTG 460
Db 301 AAAGTAAGGGCATAAAAAAGTTGATCGGAAATGTACAGATGCCAGTGGCTGCTGTAGTTG 360

Qy 461 TTATGGCTTGAATCGGGCTGATTAACCTGGAAGAGACTATTAATTCATCTTAAATATACC 520
Db 361 TTATGGCTTGAATCGGGCTGATTAACCTGGAAGAGACTATTAATTCATCTTAAATATACC 420

Qy 521 AAATATCTGTTGCGTCAAAATATCTCTTTTTCATATCCAGGATGGATCATCTCTCATG 580
Db 421 AAATATCTGTTGCGTCAAAATATCTCTTTTTCATATCCAGGATGGATCATCTCTCATG 480

Qy 581 TCAGGAAGCTTGTCTTGTAGCTATGATCAGCTGAGCTATATGACGACCTTGGATTTGAAC 640
Db 481 TTAGGAAGCTTGTCTTGTAGCTATGATCAGCTGAGCTATATGACGACCTTGGATTTGAAC 540

Qy 641 CTGTGCATCTAGGAAGACCGAGGAGCTGATGCTACTACAAAATTCAGCTCATTTACA 700
Db 541 CTGTGCATCTAGGAAGACCGAGGAGCTGATGCTACTACAAAATTCAGCTCATTTACA 600

Qy 701 AGTGGGATTCGATCGCTGTTTACAGCAATAATTTTAGCCGTGTTATCATCTAGAAG 760
Db 601 AGTGGGATTCGATCGCTGTTTACAGCAATAATTTTAGCCGTGTTATCATCTAGAAG 660

Qy 761 ATGATATGGAATAATGCCCTGATTTTGTGACTTTTGTAGGCTGGAGCTACTCTTCTTG 820
Db 661 ATGATATGGAATAATGCCCTGATTTTGTGACTTTTGTAGGCTGGAGCTACTCTTCTTG 720

```

```

Qy 821 ACAGACACAAGTCGATTATATGGCTATTCTTCTTGGAAATGCAATGACAAATGACAGTTTG 880
Db 721 ACAGACACAAGTCGATTATATGGCTATTCTTCTTGGAAATGCAATGACAAATGACAGTTTG 780

Qy 881 TCCAAGATCCTTATGCTCTTTACCGCTCAGATTTTTTCCCGGTCTTGGATGATGCTTT 940
Db 781 TCCAAGATCCTTATGCTCTTTACCGCTCAGATTTTTTCCCGGTCTTGGATGATGCTTT 840

Qy 941 CAAAATCTACTTGGGAGCAATATCTCCAAAGTGGCCAAAGGCTTACTGGGAGCTGGC 1000
Db 841 CAAAATCTACTTGGGAGCAATATCTCCAAAGTGGCCAAAGGCTTACTGGGAGCTGGC 900

Qy 1001 TAAGACTCAAAAGAGAAATCAAGAGGTCGACAATTTATTCGCCAGAAAGTTTGCAGAACAT 1060
Db 901 TAAGACTCAAAAGAGAAATCAAGAGGTCGACAATTTATTCGCCAGAAAGTTTGCAGAACAT 960

Qy 1061 ATAATTTTGGTGAGCAATGTTCTAGTTTGGGGCAGTTTTTCAAAGCAGTATCTTTGAGCCAA 1120
Db 961 ATAATTTTGGTGAGCAATGTTCTAGTTTGGGGCAGTTTTTCAAAGCAGTATCTTTGAGCCAA 1020

Qy 1121 TTAACCTAAATGATGTCCAGGTTGATTGGAAGTCAATGGAACCTTAGTTACCTTTTGAGG 1180
Db 1021 TTAACCTAAATGATGTCCAGGTTGATTGGAAGTCAATGGAACCTTAGTTACCTTTTGAGG 1080

Qy 1181 ACAATTTACGTGAACACTTTTGGTGACTTGGTTAAAGAGGCTAAGCCCATCCATGGAGCTG 1240
Db 1081 ACAATTTACGTGAACACTTTTGGTGACTTGGTTAAAGAGGCTAAGCCCATCCATGGAGCTG 1140

Qy 1241 ATGCTCTCTTGAAGCAATTTAACAATAGATGGTGATGCGTATTTCAGTACAGAGATCAAC 1300
Db 1141 ATGCTCTTGAAGCAATTTAACAATAGATGGTGATGCGTATTTCAGTACAGAGATCAAC 1200

Qy 1301 TAGACTTTGAAATATCGACGCAATTTTGGCAATTTTGAAGAAATGGAAGGATGGTGATC 1360
Db 1201 TAGACTTTGAAATATCGACGCAATTTTGGCAATTTTGAAGAAATGGAAGGATGGTGATC 1260

Qy 1361 CAGTCAGCATATAAAGGAATAGTAGTTTTCGGTACCAAAAGTCCAGACGTGATTTCC 1420
Db 1261 CAGTCAGCATATAAAGGAATAGTAGTTTTCGGTACCAAAAGTCCAGACGTGATTTCC 1320

Qy 1421 TTGTTGGCCATGATTCGCTTCAACAATCGGAATGAAGATCTTAACAAAGATATGATT 1480
Db 1321 TTGTTGGCCATGATTCGCTTCAACAATCGGAATGAAGATCTTAACAAAGATATGATT 1380

Qy 1481 GCAGGAGCCCGGCAAAAATTTTGGACTTATGGGTAGG--ATCATCGAGCTGACACAA 1538
Db 1381 ---GGAGCCCGGCAAAAATTTTGGACTTATGGGTAGGCGATCATCGAGCTGACACAA 1437

Qy 1539 ACCATGATTTTACAGTTACATACAACGTTTTTAA--TGTTATACGGAGGAGCTCACTGTT 1597
Db 1438 ACCATGATTTTACAGTTACATACAACGTTTTTAA--TGTTATATGGAGGAGCTCACTGTT 1497

Qy 1598 TAGTGTGTAAGGAGATTCGCTTCTTGTATTTGGATGAATCATCAACAAACCTATTATT 1657
Db 1498 TAGTGTGTAAGGAGATTCGCTTCTTGTATTTGGATGAATCATCAACAAACCTATTATT 1557

Qy 1658 ---TTAGTGTTCAGAACATATAAGAGAAATGTAGCCCTGTAAGACTATACATGGGACC 1714
Db 1558 AAGCCAAAGTGTTCAGAACATATAAGAGAAATGTAGCCCTGTAAGACTATACATGGGACC 1616

Qy 1715 ATCAATATCGCGG 1727
Db 1617 ATCAATATCAAG 1629

```

```

RESULT 7
NBE295993
LOCUS
DEFINITION
  Nicotiana benthamiana mRNA for N-acetylglucosaminyltransferase I
  (gntI gene).
ACCESSION
  AJ295993
VERSION
  AJ295993.1
KEYWORDS
  GntI gene; N-acetylglucosaminyltransferase I.

```

SOURCE Nicotiana benthamiana  
 ORGANISM Nicotiana benthamiana  
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
 asterids; lamnids; Solanales; Solanaceae; Nicotiana.  
 REFERENCE 1  
 AUTHORS Strasser, R., Glossl, J. and Steinkellner, H.  
 TITLE Less than 2% of GntI activity does not alter N-glycosylation in  
 Nicotiana benthamiana  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 1341)  
 AUTHORS Steinkellner, H.  
 TITLE Direct Submission  
 JOURNAL Submitted (20-SEP-2000) Steinkellner H., University of  
 Agricultural Sciences-Vienna, Center of Applied Genetics, Muthgasse  
 18, 1190 Vienna, AUSTRIA  
 FEATURES  
 source  
 1. .1341  
 /organism="Nicotiana benthamiana"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:4100"  
 1. .1341  
 /gene="gntI"  
 76. .1341  
 /gene="gntI"  
 /codon\_start=1  
 /product="N-acetylglucosaminyltransferase I"  
 /protein\_id="CAC82507.1"  
 /db\_xref="GI:27527069"  
 /db\_xref="GOA:Q8GUAS"  
 /db\_xref="InterPro:IPR004139"  
 /db\_xref="UniProt/TREMBL:Q8GUAS"  
 /translation="MRLPATQSEVADRLAAIEAENHCTSOIRLLIDHISQOGRIVA  
 LEEQKQDQRCRLVQLDLESKGIKKLGNVQMPVAAYVWACSRADYLEKTIK  
 MLKYQVSDQKRLVQLDLESKGIKKLGNVQMPVAAYVWACSRADYLEKTIK  
 KIARYKWLQDLFYKHNFSRVIILEDMEIAPDFDFFEAGATLLDRDKSIAISS  
 NDNGQMOQPYALYRSDFPFGLGWLMSKTWDELSPKWPKYMDWLRLKHNHGR  
 QFIPKVCRTNFEHGSLSGQFKOYLEPIKLANDVOVDWKSMDLSYLLEBNYVKHFG  
 NLVKAKEPIHGADAVLKAFNIDGVRIOYRQLEFEDIARQGIFFSEWKDGVPRAAKY  
 GIVFVWYQTSRRRVFLVGHDSLQQLGIET"  
 ORIGIN  
 Query Match 74.7%; Score 1297.8; DB 15; Length 1341;  
 Best Local Similarity 98.0%; Pred. No. 7,7e-307;  
 Matches 1314; Conservative 0; Mismatches 27; Indels 0; Gaps 0;  
 127 ATGAGAGGGAACAAGTTTTCCTGATTTCCGGTACCTCCTCACTTCGGCTGCTCGCC 186  
 1 ATGAGAGGGAACAAGTTTTCCTGATTTCCGGTACCTCCTCATTTTCGGCTGCTCGCC 60  
 187 TTCATCTACACAGATCGGCTTTTGGGACACAGTCAGATATGCAGATCGCCTTGC 246  
 61 TTCATCTACATACAGATCGGCTTTTTCGACACAGTCAGATATGCAGATCGCCTTGC 120  
 247 GCTGCAATTGAGCAGAAATCATTTGTAAGCCAGACAGATTCGTTATTGACCAAGATT 306  
 121 GCTGCAATTGAGCAGAAATCATTTGTAAGCCAGACAGATTCGTTATTGATCATATT 180  
 307 AGCCTGCAGCAAGGAAGATAGTTGCTCTTGAAGAACAAATGAAGCGTCAGACACAGAG 366  
 181 AGCCAGCAGCAAGGAAGATAGTTGCTCTTGAAGAACAAATGAAGCGCCAGGACGAG 240  
 367 TCCGACAAATTAGGCTCTTGTTCAGGATCTTGAAGAACAAATGAAGCGCCAGGACGAG 426  
 241 TCCGACAAATTAGGCTCTTGTTCAGGATCTTGAAGAACAAATGAAGCGCCAGGACGAG 300  
 427 GCAATGTACAGATCCAGTGCCTGTAGTTGTTATGCTTGCATCGGCTGATTAC 486  
 301 GGAATGTACAGATCCAGTGCCTGTAGTTGTTATGCTTGCATCGGCTGACTAC 360  
 487 CTGGAAGAGACTATTAAATCCATCTTTAAATAACCAATATCTGTTGCGTCAAAATATCT 546  
 361 CTGGAAGAGACTATTAAATCCATCTTTAAATAACCAATATCTGTTGCGTCAAAATATCT 420

QY 547 CTTTTCATATCCAGAGTGATCAGATCCTGATGTGAGGAAGCTTGTCTTTCAGCTATGAT 606  
 DB 421 CTTTTCATATCCAGAGTGATCAGATCCTGATGTGAGGAAGCTTGTCTTTCAGCTATGAT 480  
 QY 607 CAGTGCATATGAGCAGCACTTGGATTTTGAACCTGTGATCTGAAAGACCGGGGAG 666  
 DB 481 CAGTGCATATGAGCAGCACTTGGATTTTGAACCTGTGATCTGAAAGACCGGGGAG 540  
 QY 667 CTGATTCATATGAGCAGCACTTGGATTTTGAACCTGTGATCTGAAAGACCGGGGAG 726  
 DB 541 CTGATTCATATGAGCAGCACTTGGATTTTGAACCTGTGATCTGAAAGACCGGGGAG 600  
 QY 727 AAGCATATTTTACCGCTGTTATCATATGAGATGATATGGAATTTGCCCTTGATTTT 786  
 DB 501 AAGCATATTTTACCGCTGTTATCATATGAGATGATATGGAATTTGCCCTTGATTTT 660  
 QY 787 TTTGACATTTTTCAGGCTGAGCTACTCTTTCGACAGACAGATCGATTCGCTATT 846  
 DB 661 TTTGACATTTTTCAGGCTGAGCTACTCTTTCGACAGACAGATCGATTCGCTATT 720  
 QY 847 TCTTCTTGAATGACAAATGCAAAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 906  
 DB 721 TCTTCTTGAATGACAAATGCAAAATGCAATGCAATGCAATGCAATGCAATGCAATGCAAT 780  
 QY 907 TCAGATTTTTCAGGCTGAGCTACTCTTTCGACAGACAGATCGATTCGCTATT 966  
 DB 781 TCAGATTTTTCAGGCTGAGCTACTCTTTCGACAGACAGATCGATTCGCTATT 840  
 QY 967 CCAAAGTGGCCAAAGGCTTACTTGGGACAGATCGGCTTGAAGACAAATGCAAGATTCACAGAGGT 1026  
 DB 841 CCAAAGTGGCCAAAGGCTTACTTGGGACAGATCGGCTTGAAGACAAATGCAAGATTCACAGAGGT 900  
 QY 1027 CGACAAATTTATTCGCCCAGAGATTTTCAGAACATATATATTTTGGTGGAGCTGTTCTAGT 1086  
 DB 901 CGACAAATTTATTCGCCCAGAGATTTTCAGAACATATATATTTTGGTGGAGCTGTTCTAGT 960  
 QY 1087 TTGGGGCAGTTTTTCAAGCAGATATCTTTCAGAACATATATATTTTGGTGGAGCTGTTCTAGT 1146  
 DB 961 TTGGGGCAGTTTTTCAAGCAGATATCTTTCAGAACATATATATTTTGGTGGAGCTGTTCTAGT 1020  
 QY 1147 TGGAGTCAATGAGACCTTAGTTTACTCTTTCGAGACCAATTCGTTGAACACATCTTGGTGCAC 1206  
 DB 1021 TGGAGTCAATGAGACCTTAGTTTACTCTTTCGAGACCAATTCGTTGAACACATCTTGGTGCAC 1080  
 QY 1207 TTGGTAAAAAGGCTTAAGCCCATCCATGAGCTGATGCTGCTTGAAGACATTTAAACATA 1266  
 DB 1081 TTGGTAAAAAGGCTTAAGCCCATCCATGAGCTGATGCTGCTTGAAGACATTTAAACATA 1140  
 QY 1267 GATGCTGATGTCGCTATTTCAGTACAGAGATCAACTAGACTTTTGAAGAAATATCGACGGCAA 1326  
 DB 1141 GATGCTGATGTCGCTATTTCAGTACAGAGATCAACTAGAAATTTGAAGATATCGACGGCAA 1200  
 QY 1327 TTTGGCATTTTGAAGAAATGGAAGATGCTGTACCACTGAGCAGCATATAAAGGAATAGTA 1386  
 DB 1201 TTTGGCATTTTGAAGAAATGGAAGATGCTGTACCACTGAGCAGCATATAAAGGAATAGTA 1260  
 QY 1387 GTTTTCCGGTACCAAAAGCTCCAGAGCTGTTTCTTGTGGCATGATTTCGCTTCAACAA 1446  
 DB 1261 GTTTTCCGGTACCAAAAGCTCCAGAGCTGTTTCTTGTGGCATGATTTCGCTTCAACAA 1320  
 QY 1447 CTCGGAATTGAAGATACCTTAA 1467  
 DB 1321 CTCGGAATTGAAGATACATA 1341

RESULT 8  
 STU249878  
 LOCUS Solanum tuberosum mRNA for N-acetylglucosaminyltransferase I (GntI  
 DEFINITION gene), clone A1.  
 ACCESSION AJ249878  
 VERSION AJ249878.1 GI:18076139



Db 1358 ACTTGGAGTTGAAGTACTTACGCAAGATATGATT---GGAGCCTGAGCAACAATTTAGA 1414  
Qy 1506 CTTATTGGGTAGGATGATCGAGCTGACACTAAACCATGATTTT-----ACCAGTTAC 1558  
Db 1415 CTTATTGGTGGTACATTTTGAAGAGCTGACACGAAGATGATGACTACAGTAGCTAC 1474  
Qy 1559 ATACAACGTTTAAATGTTATACGAGGAGCTCACTGTTCTAGTGTGTAAGGATATCGGC 1618  
Db 1475 ATGCAACATTTAAATGTTAAATGGAAGNACCATGCT-----T 1513  
Qy 1619 TTCTTAGTATTGGATGATCAACACAACTATTATTTAAAGTGTTCAGAACATAAG 1678  
Db 1514 ATTGTTGAATGATGATCATCAACCATCTATTA--TTCAAAGTTTACAAACATAAG 1571  
Qy 1679 AGGAATGATGCTCTGTAAGACTA 1703  
Db 1572 AGGAATGTTGCTCTATATAAACA 1596

RESULT 9  
A95053  
LOCUS A95053 1669 bp DNA linear PAT 26-JAN-2000  
DEFINITION Sequence 1 from Patent WO9929879.  
ACCESSION A95053  
VERSION A95053.1 GI:6779205  
KEYWORDS Solanum tuberosum (potato)  
SOURCE Solanum tuberosum  
ORGANISM Solanum tuberosum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;  
asterids; lamids; Solanales; Solanaceae; Solanum.  
1 (bases 1 to 1669)  
Von,S.A.  
REFERENCE  
AUTHORS  
TITLE VEGETABLE GntI SEQUENCES AND THE USE THEREOF TO OBTAIN PLANTS WITH  
A REDUCED OR LACK OF N-ACETYLGLUCOSAMINYLTRANSFERASE I (GntI)  
ACTIVITY  
JOURNAL Patent: WO 9929879-A 1 17-JUN-1999;  
VON SCHAEWEN ANTJE (DE)  
FEATURES  
source  
1. .1669  
/organism="Solanum tuberosum"  
/mol\_type="unassigned DNA"  
/cultivar="Desiree"  
/db\_xref="taxon:4113"  
/cell\_type="BLATTZELLEN"  
/tissue\_type="MESOPHYLL"  
/clone\_lib="LAMBDA ZAP II (ECORI)"  
/dev\_stage="M: 'SINK' ORGAN"  
15. -52  
53. .1393  
/gene="CGL"  
/gene="CGL"  
53. .1393  
/EC\_number="2.4.1.101"  
/function="INITIIERT KOMPLEXE N-GLYKANE AUF SEKRET.  
GLYKOPROTEINEN"  
/standard\_name="GNTI"  
/note="ERSTE GNTI-SEQUENZ AUS KARTOFFEL (UNPUBLIZIERT)"  
/codon\_start=1  
/evidence=experimental  
/product="BETA-1,2-N-ACETYLGLUCOSAMINYLTRANSFERASE I"  
/protein\_id="CAB70462.1"  
/db\_xref="GI:6779206"

5'UTR  
Gene  
CDS  
905  
965  
891  
1025

/function="MEMBRANKER (AMINO ACID 10-29)"  
/standard\_name="MEMBRANKER EINES TYP II GOLGI-PROTEINS"  
/note="DURCH VERGLEICH MIT TIERISCHEN GNTI-SEQUENZEN  
IDENTIFIZIERT"  
/codon\_start=1  
/product="HYDROPHOBER AMINO ACIDBEREICH IN GNTI"  
/protein\_id="CAB70463.1"  
/db\_xref="GI:6779207"  
/translation="LRYLLVVAALAFIYIQMLRF"  
1394. .1655

3'UTR  
ORIGIN

Query Match 72.3%; Score 1255; DB 6; Length 1669;  
Best Local Similarity 89.1%; Pred. No. 2.3e+286;  
Matches 1412; Conservative 0; Mismatches 140; Indels 33; Gaps 4;  
Qy 126 GATGAGAGGAGGACAAAGTTTTCGTGATTTTCGGGTACTCTCTATCTTGGCTGCTGCGC 185  
Db 52 GATGAGAGGAGGACAAAGTTTTCGTGATTTTCGGGTACTCTCTATCTTGGCTGCTGCGC 111  
Qy 186 CTTTCATCTACACAGATGCGGCTTTTTCGACACAGTCAGAAATATGCGATCGCTTGC 245  
Db 112 CTTTCATCTACACAGATGCGGCTTTTTCGACACAGTCAGAAATATGCGATCGCTTGC 171  
Qy 246 TGCTGCAATTGAAGCAGAAATCATTTGTACAGCCAGACAGATGCTTTATTGACCAT 305  
Db 172 TGCTGCAATTGAAGCAGAAATCATTTGTACAGTCAGACCATGCTTTATTGCAAGAT 231  
Qy 306 TAGCTGACAGCAAGAGAGATAGTCTTCTTGAAGACAAATGAAGCTCAGGACACGGA 365  
Db 232 TAGCAGCAGCAAGAGAGATAGTCTTCTTGAAGACAAATGAAGCTCAGGACACGGA 291  
Qy 366 GTGCCGCAATTAAGGGCTCTTGTTCAGGATCTTTGAAAGTAAAGGCATATAAAGTTGAT 425  
Db 292 GTGCCGCAATTAAGGGCTCTTGTTCAGGATCTTTGAAAGTAAAGGCATATAAAGTTAAT 351  
Qy 426 CGGAAATGTACAGATGCGGCTGCTGTAGTTGTTATGCTTGCATCGGCTGATTA 485  
Db 352 CGGAGATGTGAGATGCGGCTGCTGTAGTTGTTATGCTTGCATCGGCTGATTA 411  
Qy 486 CTTGAAAGACTATTAAATCCATCTTAAATACCAATATCTGTTGCTCAAAATATCC 545  
Db 412 CTTGAGAGGAGTATTAAATCCATCTTAAATACCAATATCTGTTGCTCAAAATATCC 471  
Qy 546 TCTTTTCATATCCAGGATGATCAGCTCTGATGTCAGGAGCTTGTCTTGAAGTATGA 605  
Db 472 TCTTTTCATATCCAGGATGATCAGCTCTGATGTCAGGAGCTTGTCTTGAAGTATGA 531  
Qy 606 TCAGCTGACGTATATGACGACCTTGGATTTTGAACCTGTGATATCTGAAAGCAGGGA 665  
Db 532 TCAGCTGACGTATATGACGACCTTGGATTTTGAACCTGTGATATCTGAAAGCAGGGA 591  
Qy 666 GCTGATTCATCTACAAATTTGACGCTCATTACAGTGGGCTTGGATCAGCTGTTTGA 725  
Db 592 ACTGTTTCATCTACAAATTTGACGCTCATTACAGTGGGCTTGGATCAGCTGTTTGA 651  
Qy 726 CAAGCATATTTTAGCCGTGTTATCATATCTAGAGATGATATGAAATTTGCCCTGATTT 785  
Db 652 CAAGCATATTTTAGCCGTGTTATCATATCTAGAGATGATATGAAATTTGCTGCTGATTT 711  
Qy 786 TTTTGACTTTTTCAGGCTGAGCTTCTCTTTCAGACAGACAAAGTCGATTTATGGCTAT 845  
Db 712 TTTTGACTTTTTCAGGCTGAGCTTCTCTTTCAGACAGACAAAGTCGATTTATGGCTAT 771  
Qy 846 TTCTTCTTGGATGACAAATGACAAATGACAGTTTTCAGATCCTTATGCTCTTTACCG 905  
Db 772 TTCTTCTTGGATGACAAATGACAAATGACAGTTTTCAGATCCTTATGCTCTTTACCG 831  
Qy 906 CTCAGATTTTTCCTCGGTCTTGGATGATGCTTTTCAAAATCTACTTGGGACGAATATC 965  
Db 832 CTCAGATTTTTCCTCGGTCTTGGATGATGCTTTTCAAAATCTACTTGGGACGAATATC 891  
Qy 966 TCCAAAGTGGCCAAAGGCTTACTGGGACGACTGGCTAAAGACTCAAGAGAAATCACAGAGG 1025

Db	892	TCCTAAGTGGCCAAAGGCTTACTGGGATGCTGCTAAGGCTGAAAGAAATCACAGG	951
Qy	1026	TCGCAATTTATTCGCCAGAAAGTTTGCAGAACATATAATTTTGGTGAGCATGTTCTAG	1085
Db	952	TCGCAATTTATTCGCCAGAAAGTTTGCAGAACATATAATTTTGGTGAGCATGTTCTAG	1011
Qy	1086	TTTGGGCGAGTTTTCAGCAGTATCTTGAGCCCAATTAACATAATGATGCCAGTTGA	1145
Db	1012	TTTGGGCGAGTTTTCAGCAGTATCTTGAGCCCAATTAACATAATGATGCCAGTTGA	1071
Qy	1146	TTGGAAGTCAATGACCTTACTTACCTTTTGGAGCAATATACCTGAAACACTTTGGTGA	1205
Db	1072	TTGGAAGTCAATGACCTTACTTACCTTTTGGAGCAATATACCTGAAACACTTTGGTGA	1131
Qy	1206	CTTGGTTAAAGAGCTTAAGCCCATCCATGGAGCTGATGCTGTCTTTGAAAGCAATTAACAT	1265
Db	1132	CTTGGTTAAAGAGCTTAAGCCCATCCATGGAGCTGATGCTGTCTTTGAAAGCAATTAACAT	1191
Qy	1266	AGATGGTGAATGCGGTATTCAGTACAGAGATCAACTAGACTTTGAAAATATCGCACGGCA	1325
Db	1192	AGATGGTGAATGCGGTATTCAGTACAGAGATCAACTAGACTTTGAAAATATCGCTCGACA	1251
Qy	1326	ATTTCGGCATTTTGAAGAATGGAAGGATGGTGTACCACTGTCAGCATATATAAGGAATGT	1385
Db	1252	GTTTGGCATTTTGAAGNATGGAAGGATGGTGTACCACTGTCAGCATATATAAGGGATGT	1311
Qy	1386	AGTTTTCGGTACCAAGCTGCAGACGTGTATTCCTTTGGGCCATGATTCGCTTCAACA	1445
Db	1312	AGTTTTCGGTTCCAACATCTAGACGTGTGTCTTCTTTCCCTGATTCCTTCGACA	1371
Qy	1446	ACTCGGATTCAGATACTTAAAGATATGATTCAGAGGCCCGGCAAAATTTTGA	1505
Db	1372	ACTTCGGAGTTCAGATACTTAAAGATATGATTCAGAGGCCCGGCAAAATTTTGA	1428
Qy	1506	CTTATTGGTGGATGCGATCGAGCTGCACATAAACCATGATTTT-----ACCAGTTAC	1558
Db	1429	CTTATTGGTGGATGCGATTCATTTGAAGAGCTGCACAGAAAGTATGACTACCATGCTAC	1488
Qy	1559	ATACAACGTTTAAATGTTATACGAGGAGCTCACTGTTCTAGTGTGAAGGGATATCGC	1618
Db	1489	ATGCAACATTTTAAATGTTTAAATGGAAGAAACCACTGCT-----T	1527
Qy	1619	TTCTTAGTATTTGGATGATATCAACACAACTATTATTAAAGTGTTCAGAACATAAG	1678
Db	1528	ATTGTTGGAATGGATGATATCAACCATCTCTATTA--TTCAAAGTTTACAAACATAAAG	1585
Qy	1679	AGGAAATGAGCCCTGTAAAGACTA	1703
Db	1586	AGGAAATGTCCTTATATAAAACAA	1610
RESULT 10			
BD226904			
LOCUS	BD226904 1669 bp DNA linear PAT 17-JUL-2003		
DEFINITION	Plant GntI sequence and method of using the same for producing plant with lowered or no N-acetylglucosaminyl transferase I (GntI) activity.		
ACCESSION	BD226904		
VERSION	BD226904.1 GI:33036674		
KEYWORDS	JP 2002509696-A/1.		
SOURCE	Solanum tuberosum (potato)		
ORGANISM	Solanum tuberosum		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; asterids; lamids; Solanales; Solanaceae; Solanum.		
AUTHORS	1 (bases 1 to 1669)		
TITLE	Schaeven,A.V.		
JOURNAL	Plant GntI sequence and method of using the same for producing plant with lowered or no N-acetylglucosaminyl transferase I (GntI) activity		
	Patent: JP 2002509696-A 1 02-APR-2002;		
	ANTJE VON SCHAEWEN		

OS	Solanum tuberosum (potato)
PN	JP 2002509696-A/1
PD	02-APR-2002
PF	09-DEC-1998 JP 2000524450
PR	09-DEC-1997 DE 197 54 622.6
PI	ANTJE VON SCHAEWEN
PC	C12N15/09,A01H5/00,C07K16/40,C12N5/10,C12N9/10//C12P21/08, PC
C12N15/00	
PC	C12N5/00
CC	Strandedness: Double;
CC	Topology: Linear;
CC	/function= 'Aen codon in this context is a potential glycosylation site'
CC	/product= 'N-glycosylation consensus sequence' CC
/phenotype=	'N-glycans modulate protein properties' CC
/standard_names=	'N-glycosylation site'
CC	/label= 'pot-CHO
CC	/note= 'GntI-coding sequences from animals do not contain this feature'
CC	/codon_start= 53
CC	/function= 'initiates complex N-glycans on secretory glycoproteins'
CC	/EC number= 2.4.1.101
CC	/product= 'beta-1,2-N-acetylglucosaminyltransferase I' CC
/evidence=	EXPERIMENTAL
CC	/gene= 'cgl'
CC	/standard_name= 'gntI'
CC	/label= ORF
CC	/note= 'first gntI sequence from potato (unpublished)' CC
/function=	'membrane anchor (amino acids 10-29)' CC /product= 'hydrophobic amino acid stretch in GntI' CC /standard_name= 'membrane anchor of a type II Golgi protein' CC /note= 'identified by comparison with GntI sequences from CC animals'
CC	/function= 'used for cloning the cDNA library in Lambda ZAPII'
CC	/product= 'EcoRI/NotI-cDNA adapter'
CC	/number= 1
CC	/product= 'EcoRI/NotI-cDNA adapter'
CC	/number= 2
Key	Location/Qualifiers
FT	misc_feature 659..667
FT	CDS 53..1393
FT	5'UTR 15..52
FT	3'UTR 1394..1655
FT	CDS 80..139
FT	misc_feature 1..14
FT	misc_feature 1656..1669.
FEATURES	
source	
1..1669	
/organism= "Solanum tuberosum"	
/mol_type= "genomic DNA"	
/db_xref= "taxon:4113"	
ORIGIN	
Query Match 72.3%; Score 1255; DB 6; Length 1669;	
Best Local Similarity 89.1%; Pred. No. 2.3e-296;	
Matches 1412; Conservative 0; Mismatches 140; Indels 33; Gaps 4;	
Qy	126 GATGAGAGGGAACAAGTTTGTGTGATTTCCGGTACCTCCTCATCTTGGCTGCTGTCGC 185
Db	52 GATGAGAGGGAACAAGTTTGTGTGATTTACGGTACCTTCTCGTGTGGCTGCTCGC 111
Qy	186 CTTCTATCTACACAGATCGGGCTTTTTCGACACAGTACAGATATGCAAGATCGCCTTGC 245
Db	112 CTTCTATCTACATACAGATCGGGCTTTTTCGACACAGTACAGATATGTAAGACCGCTTGC 171
Qy	246 TGTGCAATTTGAGCAGAAATCATTTGTACAGCCAGACAGATTCGTTATTGACCAAGAT 305
Db	172 TGTGCAATTTGAGCAGAAATCATTTGTACAGCCAGACAGATTCGTTATTGACCAAGAT 231
Qy	306 TAGCCTGACAGCAAGGAAGATAGTCTCTTTGAAGAAACAAATGAAGCGTCAGGACCAAGGA 365

Db	232	TAGCCGACGAAAGGAGAGTAGCTCTTTGAAGAAACAAATGAACCATCAGGACGAGA	291
Qy	366	GTGCCGACAAATTAAGGGCTCTTGTTCAGGATCTTTGAAGTAAAGGGGATAAAAAGTTGAT	425
Db	292	GTCCGGCAATTAAGGGCTCTTGTTCAGGATCTTTGAAGTAAAGGGCATAAAAAAGTTAAT	351
Qy	426	CGGAAATGTACAGATGCCAGTGGCTGCTGTAGTTGTTATGGCTTTGCAATCGGGCTGATTA	485
Db	352	CGGAGATGTGCAGATGCCAGTGGCAGCTGTAGTTGTTATGGCTTTGCAGTCGTACTGACTA	411
Qy	486	CTTGAAAAAGACTATTAAATTCATCTTAAATACCAAATATCTGTTCCGTCAAAAATATCC	545
Db	412	CTTGGAGAGACTATTAAATTCATCTTAAATACCAAACATCTGTTGCATCAAAATATCC	471
Qy	546	TCCTTTTCATATCCGAGGATGGATCAACATCCTGATGTGAGGAAGCTTGTGTTGAGCTATGA	605
Db	472	TCCTTTTCATATCCGAGGATGGATCAAAATCCTGATGTAAAGAAAGCTTGTGAGCTATGG	531
Qy	606	TCAGCTGACGTATATGAGCAGCACTTGGATTTTGAACCTGTGCATACCTGAAAGACCGGGGA	665
Db	532	TCAGCTGACGTATATGAGCAGCACTTGGATTTATGAACCTGTGCATACCTGAAAGACCGGGGA	591
Qy	666	GCTGATTTGCATACTACAAAAATTGCACTCATTAACAAGTGGGCATTTGGATCAGCTGTTTTA	725
Db	592	ACTGTTGTCATACTACAAGATTGCAAGTCAATTACAGTGGGCATTTGGATCAGCTGTTCAC	651
Qy	726	CAAGCATAAATTTAGCCGTGTTATCATATAGAGATGATATGAAATTCGCCCTGATTT	785
Db	652	CAAGCATAAATTTAGCCGTGTTATCATATAGAGATGATATGAAATTCGCTGCTGATTT	711
Qy	786	TTTTGACTTTTTGAGGCTGAGGCTACTCTTTCTTGACAGACAGCAAGTCGATTATGGCTAT	845
Db	712	TTTTGACTATTTTTGAGGCTGAGGCTACTCTTTCTTGACAGACAGCAAGTCGATTATGGCTAT	771
Qy	846	TTCTTCTTGGAAATGACAAATGACAAATGCGAGTTTGTCCAAAGTCTCTTATGCTCTTTACCG	905
Db	772	TTCTTCTTGGAAATGACAAATGACAAAGCAGTTCGTCCAAAGTCTCTGATGCTCTTTACCG	831
Qy	906	CTCAGATTTTTTCCCGCTCTTGGATGGATGCTTTTCAAAATCTACTTTGGGACGAAATATC	965
Db	832	CTCAGATTTTTTCTTGGTCTTGGATGGATGCTTTTCAAAATCAACTTTGGTCCGAACTATC	891
Qy	966	TCCAAAGTGCCAAAGGCTTACTCGGACGACTGGCTAAGACTCAAGAGAAATCACAGAGG	1025
Db	892	TCCAAAGTGCCAAAGGCTTACTCGGATGACTGGCTAAGGCTGAAGAAATAACACAGAGG	951
Qy	1026	TCGACAAATTTATTCGCCACGAAAGTTTGCAGAACATATAAATTTTGGTGAGCATGGTCTAG	1085
Db	952	TCGACAAATTTATTCGCCACGAAAGTTTGCAGAACGCTACAATTTTGGTGAGCATGGTCTAG	1011
Qy	1086	TTTGGGGCAGTTTTTCAAGCAGTATCTTGAGCCAATTAACATAAATGATGTCCAGGTTGA	1145
Db	1012	TTTGGGGCAGTTTTTTTAAGCAGTATCTTGAGCCAATTAAGCTAATGATGTCCAGGTTGA	1071
Qy	1146	TTGGAAGTCAATTGAACCTTAGTTACTTTTTCGGAGCAATTAACGTAACAACATTTTGTGA	1205
Db	1072	TTGGAAGTCAATTGAACCTTAGTTACTTTTTCGGAGCAACATATGTGTAACAACACTTTGGCGA	1131
Qy	1206	CTTGGTTAAAAAGGCTAAGCCCATCCATGGAGCTGATGCTGTCTTGAAGCAATTAACAT	1265
Db	1132	CTTGGTTAAAAAGGCTAAGCCCATCCACGGAGCTGATGCTGTCTTGAAGCAATTAACAT	1191
Qy	1266	AGATGGTATGTGCGTATTTACGTACAGAGATCAACTAGACTTTTGAAGATATCGCACCGCA	1325
Db	1192	AGATGGTATGTGCGTATTTACGTACAGAGACCAACTAGACTTTTGAAGATATCGCTCGACA	1251
Qy	1326	ATTTGGCATTTTTTGAAGAAATGGAAGGATGGTGATCCAGCTGCAGCATATAAAGGAATAGT	1385
Db	1252	ATTTGGCATTTTTTGAAGAAATGGAAGGATGGTGATCCACGGGAGCATATAAAGGATAGT	1311
Qy	1386	AGTTTTCCGGTACCAACCGTCCAGACGTGATATCTTGTGTGGCCATGATTCGTTTCAACA	1445
Db	1312	AGTTTTCCGGTTCACAAATCTAGACGTGTGTCTTCTGTTTCCCTGATTCCTTTCGACA	1371

[illegible]







QY	606	TCAGCTGACGTATATGCAGCACTTTGGATTTTGAACTGTGTCATCTGAAAGACCAGGGGA	665
DB	566	TCACCTGACGTATATGCAGCACTTTGGATTTTGAACTGTGTCATCTGAAAGACCAGGGGA	625
QY	666	GCTGATTTGCATCTACAAAAATGCGCTCATTTACAGTGGGCATTTGGATCAGCTGTTTTA	725
DB	626	ACTGTTGTCATCTACAAGATTTGCGCTCATTTACAGTGGGCATTTGGATCAGCTGTTTTCA	685
QY	726	CAAGCATAAATTTAGCCGCTGTTATCATCTACAGAGATGATATGGAATTTGCCCTTGATTT	785
DB	686	CAAGCATAAATTTAGCCGCTGTTATCATCTACAGAGATGATATGGAATTTGCCGCTGATTT	745
QY	786	TTTTTGACTTTTTTGAGGCTGGAGCTACTCTCTTTGACAGAGACAAGTCGATTATGGCTAT	845
DB	746	TTTTTGACTACTTTTGAGGCTGGAGCTACTCTCTTTGACAGAGACAAGTCGATTATGGCTAT	805
QY	846	TTCTTTCTTGGAAATGACAAATGCGATTTGTCCAAAGATCTCTTATGCTCTTTTACCG	905
DB	806	TTCTTTCTTGGAAATGACAAATGCGATTTGTCCAAAGATCTCTTATGCTCTTTTACCG	865
QY	906	CTCAGATTTTTTCCCGGCTTTGAGTGGATGCTTTTCAAAATCTACTTTGGAGCAATTAATC	965
DB	866	CTCAGATTTTTTCCCGGCTTTGAGTGGATGCTTTTCAAAATCTACTTTGGAGCAATTAATC	925
QY	966	TCCAAAGTGGCCAAAGGCTTTACTTGGGACGACTGGCTTAAGACTCAAAGAGAAATCACAGAGG	1025
DB	926	TCCAAAGTGGCCAAAGGCTTTACTTGGGATGACTGGCTGAGGCTCAAAGAGAAATCACAGAGG	985
QY	1026	TCGACAAATTTATTTTCGCCCGAGAGTTTTCGAGAACATATATAATTTTGGTGGAGCATGGTCTTAG	1085
DB	986	TCGACAAATTTATTTTCGCCCGAGAGTTTTCGAGAACATATATAATTTTGGTGGAGCATGGTCTTAG	1045
QY	1086	TTTTGGGCGAGTTTTTCAAGCAGTATCTTGGAGCAATTTAACTAAATGATGTCACAGTTCGA	1145
DB	1046	TTTTGGGCGAGTTTTTAAAGCAGTATCTTGGAGCAATTTAACTAAATGATGTCACAGTTCGA	1105
QY	1146	TTGGAAGTCAATGGACCTTTAGTTACCTTTTTTGGAGGACAAATTAAGTGAACACATTTTGGTGA	1205
DB	1106	TTGGAAGTCAATGGACCTTAAGTTACCTTTTTTGGAGGACAACTATGTAACACATTTTGGAGA	1165
QY	1206	CTTGGTTAAAAAGGCTAAGCCCATCCCATGAGAGCTGATGCTGTTGAAAGCATTTTAACAT	1265
DB	1166	CTTGGTTAAAAAGGCTAAGCCCATCCCATGAGAGCTGATGCTGTTTGAAGACATTTTAACAT	1225
QY	1266	AGATGGTGATGTCGGTATTCAGTACAGAGATCAACTAGACTTTTGAAATATCCACGGCA	1325
DB	1226	AGATGGTGATGTCGGTATTCAGTACAGAGATCAACTAGACTTTTGAAAGCATTTTAACAT	1285
QY	1326	ATTTGGCATTTTTGAAGAAATGGAAGGATGGTGATCCACCTGACAGCATATAAAGGAATAGT	1385
DB	1286	GTTTGGAAATTTTGAAGAAATGGAAGGATGGTGATCCACCTGACAGCATATAAAGGAATAGT	1345
QY	1386	AGTTTTCCGGTACCAACAGTCCAGACGTGATTTCTCTTTGGTGGCCCATGATTCGCTTCACA	1445
DB	1346	AGTTTTCCGGTACCAACATCTAGACGTGTTGTTCTTTATTTGCCCTGATTTCTCTTCAACA	1405
QY	1446	ACTCGGAATTCGAAGATCTTAAACAAGATATGATGTCAGGAGCCCGGGCAAAATTTTTGA	1505
DB	1406	ACTTGGAGTTGAAGATCTTAGCAAGATATGATTTTGGAGCGTGAGCAACAAATTTAGA	1465
QY	1506	CTTATTTGGGTAGGATGCAAT-----CGAGCTGACACTAAACCATGATTTT--TACCAGTTAC	1565
DB	1463	CTTATTTGGCAGGATACATTTGAAGAGCTGGCAGCAAGAGCTTGATTAGCAGTAGCTGC	1522
QY	1559	ATACAAAGTTTTTAATGTTATACGGAGAGCTCAGCTGTTCTAGTGTTCGAAGGGAT	1612
DB	1523	ATACAAATTTTAATGTTATGTTAAATGGAAGGAACCTCACTGCCTACTGTTGGATGGAT	1576
RESULT 13			
STU249879			
LOCUS			
		1691 bp	mRNA linear
		STU249879	PLN 15-APR-2005

Db 184 GTCAGGACCAGGAGTCCGGCAATTAAGGGCTCTTGTTCCAGGATCTTTGAAAGTAAGGGCA 243  
Qy 413 TAAAAAGTTGATCGGAAATGTACAGATGCGAGTGGCTGCTGTAGTTGTTATGCTTGCA 472  
Db 244 TAAAAAGTTAATCGGAGATGTGCAGATGCGAGTGGCGGCTGTAGTTGTTATGCTTGCA 303  
Qy 473 ATGGGCTGATTACTCGTGAAGAACTATTAAATCCATCTTTAAATACCAATATCTGTG 532  
Db 304 GTCTGTCTGACTCTCGGAGAAGACTATTAAATCCATCTTTAAATACCAATATCTGTG 363  
Qy 533 CGTCAAAATATCTCTTTTTCATATCCAGGATGGATCACATCCTGTATGTCAGGAAGCTTG 592  
Db 364 CATCAAAATATCTCTTTTTCATATCCAGGATGGATCAATCTCTGTATGTTAGAAAGCTTG 423  
Qy 593 CTTTGACTATGATCAGCTGACGTATATGAGAGCACTTGGATTTTGAACCTGTGCATCTG 652  
Db 424 CTTTGACTATGATCAGCTGACGTATATGAGAGCACTTGGATTTTGAACCTGTGCATCTG 483  
Qy 653 AAAGACGAGGGAGCTGATTGCTACTACTACAAATTTGCAAGTTCATTAACAGTGGGCAATGG 712  
Db 484 AAAGACGAGGGAACTGGTTGCATACTACAGATTTGCAAGTTCATTAACAGTGGGCAATGG 543  
Qy 713 ATCAGCTGTTTAAACAGCAATAATTTAGCCGTGTTATCATACTAGAGATGATATGAAA 772  
Db 544 ATCAGCTGTTTCAACAGCAATAATTTAGCCGTGTTATCATACTAGAGATGATATGAAA 603  
Qy 773 TTGCCCTGATTTTTTGGACTTTTTTGGAGTGAGGCTACTCTTCTTGAAGAGCAAGT 832  
Db 604 TTGCTGCTGATTTTTTGGACTTTTTTGGAGTGAGGCTACTCTTCTTGAAGAGCAAGT 663  
Qy 833 CGATTATGGCTATTCTCTTGAAGTACAAATGCAAGTTCAGTCTTCTTGAAGAGCAAGT 892  
Db 664 CGATTATGGCTATTCTCTTGAAGTACAAATGCAAGTTCAGTCTTCTTGAAGAGCAAGT 723  
Qy 893 ATGCTCTTTACCGCTCAGATTTTTTCCCGCTCTTGATGATGCTTTTCAAAATCTACTT 952  
Db 724 ATGCTCTTTACCGCTCAGACTTTTTTCCCGCTCTTGATGATGCTTTTCAAAATCAACTT 783  
Qy 953 GGGACGAATTTATCTCAAAAGTGGCCAAAGGTTTACTGGGACGACTGGCTAAGACTCAAG 1012  
Db 784 GGTCCGAATCTATCTCAAAAGTGGCCAAAGGTTTACTGGGACGACTGGCTAAGGCTCAAG 843  
Qy 1013 AGAATCACAGAGTTCGCAATTTATTCGCCAGAGTTTCGAGAACATATAATTTGGTG 1072  
Db 844 AAAATCACAGAGTTCGCAATTTATTCGCCAGAGTTTCGAGAACATATAATTTGGTG 903  
Qy 1073 AGCATGTTCTAGTTTGGGCGAGTTTTTCAAGCAGTATCTTGAGCCCAATTAACCTAAATG 1132  
Db 904 AGCATGTTCTAGTTTGGGCGAGTTTTTCAAGCAGTATCTTGAGCCCAATTAAGCTAAATG 963  
Qy 1133 ATGTCAGGTTGATTGGAAGTCAATGGAACCTTAGTTACCTTTTGGAGGACAATTAAGCTGA 1192  
Db 964 ATGTCAGGTTGATTGGAAGTCAATGGAACCTTAGTTACCTTTTGGAGGACAATTAAGCTGA 1023  
Qy 1193 AACACTTTGTGTGCTTGTAAAGGCTAAGCCATCCAGCTGATGCTGCTGTTGA 1252  
Db 1024 AACACTTTGTGTGCTTGTAAAGGCTAAGCCATCCAGCTGATGCTGCTGTTGA 1083  
Qy 1253 AAGCATTTAAATAGATGTTGATGCTGCTATTCAGTACAGAGATCAACTAGACTTTGAAA 1312  
Db 1084 AAGCATTTAAATAGATGTTGATGCTGCTATTCAGTACAGAGATCAACTAGACTTTGAAA 1143  
Qy 1313 ATATCGCAGCGCAATTTGGCAATTTTGAAGATGGAAGGATGGTGATCAACGTCAGCAT 1372  
Db 1144 ATATCGCTCGCGCAATTTGGCAATTTTGAAGATGGAAGGATGGTGATCAACGTCAGCAT 1203  
Qy 1373 ATAAAGGAATAGTAGTTTTCGGTACCAGCTCCAGCTGATCTCTTGTGGCCATG 1432  
Db 1204 ATAAAGGAATAGTAGTTTTCGGTACCAGCTCCAGCTGATCTCTTGTGGCCATG 1263  
Qy 1433 ATTGCTCTCAACACTCGGAATTTGAAGATATCTTAACAAAGATATGATTCAGGAGGCCGG 1492  
Db 1264 ATTCTCTTCGACAACTTGGAGTTGAAGATACGTAGCGAAGATATGATT--TGTGCGCTGA 1321

Qy 1493 GCAAAATTTTGCATTTATTTGGGTAGCATGCAT-----CGAGCTGACACTAAA--CCATGA 1545  
Db 1322 GCAACCTTTTAGACTTTATTTGGTAGGATACATTTGAAGGAGCTGACACGAAAAGCATGAC 1381  
Qy 1546 TTTTACCAAGTTACATCAACAGTTTAAATGTTATACGAGGAGCTCACTGTTCTAGTGTG 1605  
Db 1382 TACCAGTACTACATGCAACATTTTAAATGTTATGAAGGAACCTCACTGCCTACTGTTGG 1441  
Qy 1606 AAGGAT 1612  
Db 1442 AATGAT 1448

RESULT 14  
STU249880

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## FEATURES

## source

## gene

## CDS

## ORIGIN

## Query Match

## Best Local Similarity

## Score 1078;

## DB 15;

## Length 1499;

## Pred. No. 5.3e-253;

## 62.1%;

## 87.5%;

## 1499;

## 1499;

## 1499;

## 1499;

## 1499;

## 1499;

## 1499;

## 1499;

## 1499;

## 1499;

## 1499;

## 1499;

## 1499;

## 1499;

## 1499;

## 1499;

## 1499;

## 1499;

## 1499;

## 1499;

## 1499;

## 1499;

## 1499;

## 1499;

## 1499;

## 1499;

## 1499;

## 1499;

## 1499;

## 1499;

## 1499;

Matches 1237; Conservative 0; Mismatches 145; Indels 32; Gaps 4;	
Qy	330 TGCTCTTGAAGAACAAATGAAGCTCAGGACCAGAGTGGCGACAATTAAGGGCTCTTGT
Db	21 TGTCTCCGGTAAACAAATGAAGCTCAGGACCAGAGTGGCGCAAATTAAGGGCTCTTGT
Qy	390 TCAGGATCTTGAAGTAAAGGCATAAAAAGTTGATCGGAAATGTACAGATGCCAGTGGC
Db	81 TCAGGATCTTGAAGTAAAGGCATAAAAAGTTAATCGGAGATGTCAGATGCCAGTGGC
Qy	450 TGCTGTAGTTATGCTTGCATTCGGGCTGATTAACCTGGAAAAAGCATTAATATCAT
Db	141 AGCTGTAGTTATGCTTGCATTCGGGCTGATTAACCTGGGAGAGCATTAATATCAT
Qy	510 CTTAAATACCAATATCTGTGGTCAAAATATCCTTTTCATATCCAGGATGGATC
Db	201 CTTAAATACCAATATCTGTGGTCAAAATATCCTTTTCATATCCAGGATGGATC
Qy	570 ACATCCTGATGTCAGGAAGCTTGTCTTGAGCTATGATCAGCTGACGTATATGACGACTT
Db	261 AATCCTGATGTAAGAAGCTTGTCTTGAGCTATGATCAGCTGACGTATATGACGACTT
Qy	630 GGATTTTGAACCTGTGCATCTGAAAGACAGGAGGAGCTGATGCATCTACATAAAATGC
Db	321 GGATTTTGAACCTGTGCATCTGAAAGACAGGAGGAACTGGTTGCATCTACATAAAATGC
Qy	690 ACCTCATACAGTGGGCAATTTGGATCAGCTGTTTACAAAGCATATTTAGCCGTGTAT
Db	381 ACCTCATACAGTGGGCAATTTGGATCAGCTGTTTACAAAGCATATTTAGCCGTGTAT
Qy	750 CATACTAGAAGATGATATGGAATATGCCCTCGATTTTTCGACTTTTTCAGGCTGGAGC
Db	441 CATACTAGAAGATGATATGGAATATGCTGCTGATTTTTCGACTTTTTCAGGCTGGAGC
Qy	810 TACTCTTCTGACAGAGACAAGTCGATATGCTATTTCTTCTGGAATGACAATGAGACA
Db	501 TACTCTTCTGACAGAGACAAGTCGATATGCTATTTCTTCTGGAATGACAATGAGACA
Qy	870 AATGAGTTTGTCCAGATCTTATGCTCTTTACCGCTCAGATTTTTCGCGTCTTGG
Db	561 AAGGAGTTTGTCCAGATCTTATGCTCTTTACCGCTCAGATTTTTCGCGTCTTGG
Qy	930 ATGATGCTTTCAAAATCTACTTTGGGAGCAATATCTCAAAAGTGGCCAAAGGCTTACTG
Db	621 ATGATGCTTTCAAAATCAACTTGTTCGAACTATCTCAAAAGTGGCCAAAGGCTTACTG
Qy	990 GGACGACTGGCTAAGACTCAAGAGAAATCAGAGGTCGACAATTTATTCGCCAGAAAT
Db	681 GGATGACTGGCTAAGGCTGAAAGAAATCAGAGGTCGACAATTTATTCACCCAGAAAT
Qy	1050 TTGCAGAACATATAATTTTGTGAGCATGTTCTAGTTTGGGCAAGTTTTCAGCAGTA
Db	741 TTGCAGAACGTCATAATTTTGTGAGCATGTTCTAGTTTGGGCAAGTTTTCAGCAGTA
Qy	1110 TCTTGAGCAATTAACATAATGATGTCAGGTCGATTTGGAAGTCAATGAGCTTTAGTTA
Db	801 TCTTGAGCAATTAAGCTAAATGATGTCAGGTCGATTTGGAAGTCAATGAGCTTTAGTTA
Qy	1170 CTTTGTGGAGCAATTAAGTGAACACACTTTTGGTACTTTGGTTAAAAAGGCTTAAGCCAT
Db	861 CTTTGTGGAGCAACTATGTAACACTTTTGGGCACTTTGTTAAAAAGGCTTAAGCCAT
Qy	1230 CCAATGAGCTGATGCTGCTTGAAGCAATTAACATAGATGATGTCGATTTACGTA
Db	921 CCACGGAGCTGATGCTGTTTGAAGCAATTTAAACATAGATGATGTCGATTTACGTA
Qy	1290 CAGAGATCAACTAGACTTTGAAATATCGACGGCAATTTGCAATTTTGAAGATGAA
Db	981 CAGAGCAACTAGACTTTGAAAGCTATCGCTCGACAGTTTGGCATTTTGAAGATGAA
Qy	1350 GGATCGTGTACCACTGACGATATAAGAGGAATAGTATTTTCCGGTACCAACGTCAG
Db	1041 GGATCGTGTACCACTGACGATATAAGGGATAGTATTTTCCGGTTCACCAACATCTAG

Qy	1410 ACGTGATATCTTGTGGCCATGATTCGCTTCAACAACTCGAATTGAAGATACTTAACA
Db	1101 ACGTGATATCTTGTGGCCCTGATTCCTTCGACAACTTGGAGTTGAAGATACTTAGCG
Qy	1470 AAGATATGATGACAGAGCCCGGCAAAATTTTGAATTTGGTGGTGGATGATCGAGC
Db	1161 AAGATATGATTTTGGAGCCTGAGCAAAATTTAGACTTTATTTGGTGGTGGATCGAA
Qy	1530 TGACACTAAACCATGATTTT-----ACCAGTTACATACAAAGCTTTTAAATGTTATACGG
Db	1218 AGAGCTGACAGAAAGTATGACTACCAAGTAGCTACATGCAACATTTTAAATGTTAATGGA
Qy	1583 AGGAGCTCACTGTTCTAGTGTGGAAGGATATCGGCTTCTTAGTATTGGAATGAATCATCA
Db	1278 AGGAACCCACTGCT-----TATTGTTGGAATGGATGAATCATCA
Qy	1643 ACACAACTTATTTTAAAGTTTTCAGAACATAAAGAGGAAATGTAGCCCTGTAAAGACT
Db	1317 CCACATCTTATTA-TTCAAGTTTTCACAAACATAAAGAGGAAATGTAGCCCTTATAACACA
Qy	1703 ATACATGGACCATCAATAATCGCGCGCGCAATT
Db	1376 ATGTTTTTGTCTTAAAGAGAACGTTACGATTT

RESULT 15	
BT000334	1625 bp mRNA linear PLN 19-SEP-2002
LOCUS	Arabidopsis thaliana glycosyltransferase-like protein (At4g38240)
DEFINITION	mRNA, complete cds.
ACCESSION	BT000334.1
VERSION	BT000334.1 GI:23198251
KEYWORDS	FLI CDNA.
SOURCE	Arabidopsis thaliana (thale cress)
ORGANISM	Arabidopsis thaliana
REFERENCE	1 (bases 1 to 1625)
AUTHORS	Nguyen,M., Karlin-Neumann,G., Southwick,A., Tripp,M., Miranda,M., Palm,C.J., Bowser,L., Jones,T., Banh,J., Carninci,P., Chen,H., Cheuk,R., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A., Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K., Sakano,H., Sakurai,T., Satou,M., Seki,M., Shinn,P., Yamada,K., Shinozaki,K., Ecker,J.J., Theologis,A. and Davis,R.W.
TITLE	Direct Submission
JOURNAL	Submitted (19-SEP-2002) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
COMMENT	e-mail for correspondence: arab@sequence.stanford.edu

RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of AFL cDNAs (RAFL cDNA : 'RIKEN Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J., Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.

The Salk, Stanford, PGSC (SSP) Consortium members constructed and sequenced the pENTR (ORF) clones using the RAFL cDNAs: Nguyen,M., Southwick,A., Tripp,M., Palm,C.J., Jones,T., Wu,T., Chen,H., Cheuk,R., Chan,M.M., Chang,C.H., Dale,J.M., Deng,J.M., Hsuan,V.W., Lee,J.M., Kim,C.J., Quach,H.L., Shinn,P., Tang,C.C., Toroumi,M., Wallender,E.K., Wong,C., Wu,H.C., Yamada,K., Yu,G., Yuan,S., Ecker,J.J., Theologis,A. and Davis,R.W.

Nguyen,M., (SSP/Stanford) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Davis,R.W. (SSP/Stanford) contributed equally to this work as PIs.

FEATURES	Location/Qualifiers
source	1..1625
	/organism="Arabidopsis thaliana"
	/mol_type="mRNA"

gene	/db_xref="taxon:3702"	
	/chromosome="4"	
CDS	/clones="U25541"	
	/ecotype="Columbia"	
	/note="This clone is in pENTR/SD-dTopo This is a cloned PCR product using RIKEN clone RAFL09-98-A11 (AY099838) as a template"	
	1..1625	
	/gene="At4g38240"	
	/note="synonym: F20D10.360"	
	1..1335	
	/gene="At4g38240"	
	/codon_start=1	
	/product="glycosyltransferase-like protein"	
	/protein_id="AA15653.1"	
	/db_xref="GI:23198252"	
	/translation="MARISSDLRLFLIPAAFMFIYIQMLRFQTSQVADRLSSAIBSE NHCTSMRGILIDEVSIKSRIVALEDMKQDELVOLKDIOTFEKKGIKLTQGGQ MPVAADVVMACSRADYILERTVKSVLTYQTVPASKYPLFISODGSDQAKVSKLSYNOL TYMHQLDFEPVTPERBELTAYYKIAHRYKWLDDLPYKHKFSRVILLEDMEIADPF FDYFEAASLMDRDKTIMAASSMNDNGKQFVHDPIYALYRSDFPGLGMLKSTWDE LSPWKAYADWDMLRLKENHKGRQFIRPEVCRTYNFGEHSSLSGFQFSYLEPIKUND VTVDKADLGLYLEGNYTKYFSLVRQARPIQGSIDLVLKAQNIKDDVRIRYKQDVEF ERIAGEFGIFEWKDGVPTAYKGVVVFRIQTTRRVFLVGPDSVMQLGRNS"	
	ORIGIN	
	Query Match 44.4%; Score 772; DB 15; Length 1625;	
	Best Local Similarity 73.3%; Pred. No. 4.7e-178;	
	Matches 988; Conservative 0; Mismatches 360; Indels 0; Gaps 0;	
Qy	139	AAGTTTCTGTGATTCGGGTACTCTCTCATCTTGGCTGTGCTGCGCTTCATCTACACA 198
Db	7	AGGATCTCGTGACTTGAGATTTCTTCTATCCCGGCGAGCTTTCATGTTCAITCATC 66
Qy	199	CAGATGGCGCTTTTGGCGACACAGTCAGAAATATCAGATCGCCTTGCTGTGCAATTTGAA 258
Db	67	CAGATGAGGCTTTTCCAGACGCATCAGATGATCGAGATCGCCTCAGTTCGGCTATCGAA 126
Qy	259	GCAGAAATCTTTGTACAGCCAGACAGATGTTCTTATTCACAGATTAGCCTGCAGCAA 318
Db	127	TCTGAGAACCAATTGCACCTAGTCAAATCGAGGCCCTCATAGATGAAGTTAGCATCAAAACAG 186
Qy	319	GGAGATAGTGTCTTGAAGACAAATGAAGCGTCAGGACCGAGGAGTGCCGCAATTA 378
Db	187	TCGGCGATGTGTGCCCTCGAGATATGAAGAACCGCCAGGACGAAGAACTTTGTCAGCTT 246
Qy	379	AGGCTCTTGTTCAGGATCTTGAAGTAAGGCGCAAAAAAGTTGATCGGAAATGTACAG 438
Db	247	AAGGATCTAATCCAGACGTTTGAANAAGAGATAGCNAAACTCACTCAAGTTGGACAG 306
Qy	439	ATGCCAGTGGCTGCTAGTGTGTTATGGCTTGGCAATCGGGCTGATTAACCTGGAAAAAGACT 498
Db	307	ATGCCGTGTGGCTGCTAGTGTGATGCGCTGCGAGTGGTGCAGACTATCTTGAAGGAGCT 366
Qy	499	ATTAAATCCATCTTAATACCAATATCTGTTCGGTCAGAAATATCTCTTTTCATATCC 558
Db	367	GTTAAATCAGTTTAAACATATCAAACTCCCGTGTCTTCAAAAATATCTCTATTTATATCT 426
Qy	559	CAGGATGGATCACATCTCGATGTGAGGAAGCTTGCTTTGAGCTATGATCAGCTGACGTAT 618
Db	427	CAGGATGGATCTGATCAAGCTGTCAAGAGCAAGTCAITTAGCTATATCAATTAAACATAT 486
Qy	619	ATGAGCACTTTGGATTTTGAACCTGTGCAATCTGAAAGACAGGGGAGCTGATTGCAATAC 678
Db	487	ATGAGCACTTTGGATTTTGAACCAAGTGTGCTCAAAAGGCCCTGGCGAACTGACTGCGTAC 546
Qy	679	TACAAATTCGACCTCATTAAGTGGCAATTTGATCAGCTGATGTTTACAGCATAAATTTT 738
Db	547	TACAAGATTGCACTGCTACCTCAAGTGGGCACTGGACCAAGTTGTTTACAAACACAAATTT 606
Qy	739	AGCGGTGTTATCATCATAGAGATGATGAAATTTGCCCTGATTTTGTGACTTTTTT 798
Db	607	AGTCGAGTGATTATACATAGAGATGATGAAATTTGCTCCAGACTTCTTTTGATTACTTTT 666

Search completed: December 14, 2005, 11:16:02  
Job time : 8633 secs

Qy	799	GAGGCTGGAGTACTCTCTTTCACAGAGACAAGTCGATTATGGCTATTCTTCTTCTGGAAT 858
Db	667	GAGGCTGCAGCTAGTCTCATGGATAGGGATAAAACCAATTATGGCTGCTTCATCATGGAAT 726
Qy	859	GACAATGGACAATGCAAGTTTGTCCAAAGATCCTTATGCTCTTTACCGCTCAGATTTTTTT 918
Db	727	GATNATGGACAGNAGCAGTTTGTGCATGATCCTATGCGCTATACCGATCAGATTTTTTT 786
Qy	919	CCCGTCTTTGGATGGATGCTTTCAAAATCTACTTTGGGACGAATATCTCCAAAGTGGCCA 978
Db	787	CTTGGCTTTGGGTGGATGCTCAAGAGATCGACTTGGGATGAGTTATCACCAAAAGTGGCCA 846
Qy	979	AAGGCTTACTGGGACCACTGGCTAAGACTCAAGAGAAATCACAGAGTGCAGACATTTATT 1038
Db	847	AAGGCTTACTGGGATGATTGGCTGAGACTAAAGGAAACCATAAAGGCCGCCCAATTCATT 906
Qy	1039	CGCCCAAGAGTTTGCAGAAACATATAATTTTGTGAGCATGGTTCTAGTTTGGGGCAGTTT 1098
Db	907	CGACCGGAAGTCTGTAGAACATACAAATTTTGGTGAACATGGGTCTAGTTTGGGACAGTTT 966
Qy	1099	TTCAAGCAGTATCTTGAGGCCAAATTAACCTAAATGATGTCAGGTTGATTTGGAAGTCAATG 1158
Db	967	TTCAGTCAGTATCTGGAAACCTATAAAGCTAAACGATGTGACGTTGACTGGAAGCAAG 1026
Qy	1159	GACCTTAGTTACCTTTTGGAGGACAATTAGCTGAAACACTTTTGGTGAATTTGTTAAAAAG 1218
Db	1027	GACCTGGGATACCTGACAGAGGGGAAACTATACCAAGTACTTTTCTGGCTTAGTGAGACAA 1086
Qy	1219	GCTAAGCCCATCCATGGAGCTGATGCTGCTTTGAAAGCATTTTAAACATAGATGCTGATGTG 1278
Db	1087	GCAGGACCAATTCAGGTTCTGACCTTGCTTTAAAGGCTCAAAACATAAAGATGATGTT 1146
Qy	1279	CGTATTCACTACAGAGATCAACTAGACTTTTGAANAATATCGCACGCAATTTGGCAATTTT 1338
Db	1147	CGTATCCGGTATAAAGACCAAGTAGAGTTTGAACGCAATTCGAGGGGAAATTTGGTATATTT 1206
Qy	1339	GAAGATGGAAGGATGGTGTACCACTGCGAGCATATAAAGGAATAGTATTTTCCGGTAC 1398
Db	1207	GAAGATGGAAGGATGGTGTGCTTCGAAACAGCATATAAAGGAGTAGTGGTGTTCGGAATC 1266
Qy	1399	CAAACTGTCAGAGCTGTATTCTTGTGTCATGATTTCGCTTCAACAACTCCGGAATTTGAA 1458
Db	1267	CAGACAAACAGAGCTGTATTCTGTTGGGCGAGATTCGTATATGAGCTTGGAAATTCGA 1326
Qy	1459	GATACCTTAAACAAAGATATGATTCAGGA 1486
Db	1327	AATTCCTGATGCAAAACATATGAAGGA 1354

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2005, 08:25:25 ; Search time 339 Seconds

(without alignments)  
9108.037 Million cell updates/sec

Title: US-10-615-144-3

Perfect score: 1737

Sequence: 1 gaattcgccgcattga.....ataatcgccgcgaattc 1737

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA.\*

- 1: /cgn2\_6/ptodata/1/ina/1 COMB.seq.\*
- 2: /cgn2\_6/ptodata/1/ina/5 COMB.seq.\*
- 3: /cgn2\_6/ptodata/1/ina/6A COMB.seq.\*
- 4: /cgn2\_6/ptodata/1/ina/6B COMB.seq.\*
- 5: /cgn2\_6/ptodata/1/ina/H COMB.seq.\*
- 6: /cgn2\_6/ptodata/1/ina/pCTUS COMB.seq.\*
- 7: /cgn2\_6/ptodata/1/ina/pp COMB.seq.\*
- 8: /cgn2\_6/ptodata/1/ina/RE COMB.seq.\*
- 9: /cgn2\_6/ptodata/1/ina/backfiles.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1737	100.0	1737	US-09-591-466C-3	Sequence 3, Appl
2	1255	72.3	1669	US-09-591-466C-1	Sequence 1, Appl
3	767.2	44.2	1854	US-09-591-466C-5	Sequence 5, Appl
4	45.4	2.6	1141	US-09-806-708B-22	Sequence 22, Appl
5	45.4	2.6	1141	US-09-806-708B-22	Sequence 22, Appl
6	40	2.3	778	US-09-270-767-9920	Sequence 9920, Ap
7	38.8	2.2	2793	US-09-134-001C-2477	Sequence 2477, Ap
8	38.8	2.2	2976	US-09-386-962C-7	Sequence 7, Appl
9	38.8	2.2	2976	US-09-386-959-7	Sequence 7, Appl
10	38.2	2.2	942	US-09-248-796A-2786	Sequence 2786, Ap
11	36.8	2.1	32172	US-09-949-016-16440	Sequence 16440, A
12	36.6	2.1	2353	US-09-620-312D-194	Sequence 194, App
13	36.6	2.1	2725	US-10-012-231A-87	Sequence 87, Appl
14	36.6	2.1	2725	US-10-015-389A-87	Sequence 87, Appl
15	36.6	2.1	2725	US-10-006-768A-87	Sequence 87, Appl
16	36.6	2.1	2725	US-10-015-671A-87	Sequence 87, Appl
17	36.6	2.1	2725	US-10-015-393A-87	Sequence 87, Appl
18	36.6	2.1	2725	US-10-011-833A-87	Sequence 87, Appl
19	36.6	2.1	2725	US-10-006-041A-87	Sequence 87, Appl
20	36.6	2.1	2725	US-10-012-064A-87	Sequence 25, Appl
21	36.6	2.1	2731	US-09-786-240-25	Sequence 14000, A
22	36.4	2.1	154626	US-09-949-016-14000	Sequence 313, App
23	36	2.1	858	US-09-107-532A-313	Sequence 16420, A
24	36	2.1	234884	US-09-949-016-16420	Sequence 16420, A

25	35.8	2.1	1497	3	US-09-270-767-11259	Sequence 11259, A
C 26	35.8	2.1	2934	2	US-08-198-446B-16	Sequence 16, Appl
C 27	35.8	2.1	2934	2	US-08-870-693-16	Sequence 16, Appl
C 28	35.8	2.1	237241	3	US-09-949-016-16101	Sequence 16101, A
29	35.6	2.0	832	3	US-09-621-976-2813	Sequence 2813, Ap
30	35.4	2.0	2328	3	US-09-801-198-49	Sequence 49, Appl
31	35.4	2.0	151256	3	US-09-949-016-12674	Sequence 12674, A
32	35.4	2.0	151261	3	US-09-949-016-13242	Sequence 13242, A
C 33	34.8	2.0	933	3	US-09-248-796A-13724	Sequence 13724, A
C 34	34.8	2.0	19227	3	US-09-090-793-13	Sequence 13, Appl
C 35	34.8	2.0	19227	3	US-09-231-899-13	Sequence 13, Appl
C 36	34.8	2.0	40138	3	US-09-090-793-12	Sequence 12, Appl
C 37	34.8	2.0	40138	3	US-09-231-899-12	Sequence 12, Appl
C 38	34.8	2.0	154915	3	US-09-949-016-15584	Sequence 15584, A
C 39	34.6	2.0	1092	3	US-09-498-520A-35	Sequence 35, Appl
40	34.6	2.0	1609	3	US-09-312-285-1	Sequence 1, Appl
41	34.6	2.0	1609	3	US-09-728-764-1	Sequence 1, Appl
C 42	34.4	2.0	1392	3	US-09-248-796A-854	Sequence 854, App
C 43	34.4	2.0	670689	3	US-09-949-016-12505	Sequence 12505, A
C 44	34.4	2.0	670690	3	US-09-949-016-14207	Sequence 14207, A
45	34.2	2.0	14273	3	US-08-961-527-40	Sequence 40, Appl

ALIGNMENTS

RESULT 1  
US-09-591-466C-3  
; Sequence 3, Application US/09591466C  
; Patent No. 6653459  
; GENERAL INFORMATION:  
; APPLICANT: Von Schaeuwen, Antje  
; TITLE OF INVENTION: Plant GntI Sequences and the Use Thereof for the Production  
; OF PLANTS HAVING REDUCED OR LACKING N-acetyl Glucosaminyl  
; TRANSFERASE I (GntI) Activity  
; TITLE OF INVENTION: Transferrase I (GntI) Activity  
; FILE REFERENCE: 032266-003  
; CURRENT APPLICATION NUMBER: US/09/591.466C  
; CURRENT FILING DATE: 2000-06-09  
; PRIOR APPLICATION NUMBER: EP 98/08001  
; PRIOR FILING DATE: 1998-09-12  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 1737  
; TYPE: DNA  
; ORGANISM: Nicotiana tabacum  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (733)...(741)  
; OTHER INFORMATION: function: Asn codon in this context is a potential  
; glycosylation site;  
; OTHER INFORMATION: glycosylation site;  
; OTHER INFORMATION: product: N-glycosylation consensus sequence;  
; OTHER INFORMATION: phenotype: N-glycans modulate protein properties;  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (733)...(741)  
; OTHER INFORMATION: standard name: N-glycosylation site;  
; OTHER INFORMATION: label: pot-CHO;  
; OTHER INFORMATION: note: gntI sequences from animals do not contain  
; this feature.  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (127)...(1467)  
; OTHER INFORMATION: codon\_start: 127;  
; OTHER INFORMATION: function: initiates complex N-glycans on secretory  
; glycoproteins;  
; OTHER INFORMATION: glycoproteins;  
; OTHER INFORMATION: EC\_number: 2.4.1.101;  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (127)...(1467)  
; OTHER INFORMATION: product: beta-1,2-N-acetylglucosaminyltransferase I;  
; OTHER INFORMATION: evidence: EXPERIMENTAL;  
; OTHER INFORMATION: gene: cgl;

= parent of restant  
appl











Qy	199	CAGATCGGCTTTTTCGACACACAGTCAGATATGCAGATCGCTCTGCTGCAATTTGAA	258
Db	201	CAGATGAGGCTTTTTCAGACCGCAATCACAGTATGCAGATCGCCTCAGTTCCGCTATCGAA	260
Qy	259	GCAGAAATCATTTGTATCAAGCCAGACAGAGTTCTTTATTGACCAGATTAGCTTCGACGAA	318
Db	261	TCTGAGAACCATTTGCATCTAGTCAAATTCGAGGGCTCATAGATGAGTTAGCATCAAAACAG	320
Qy	319	GGAGAAATAGTTGCTCTTTGAAGAACAAATGAAGCGTCGAGCACGAGAGTGCOCACAATTA	378
Db	321	TCGCGGATTTGCCCTCGAAGATATGAAGAACCGCCAGGACGAGAACTTGTGCGAGCTT	380
Qy	379	AGGCTCTTGTTCAGGATCTTTGAAGTAAGGGCATATAAAGTTGATCGGAAATGTACAG	438
Db	381	AAGGATCTAAATCCAGACGTTTGAATAAAGAAAGTAATACRAAACTCACTCAAGGTGACAG	440
Qy	439	ATCCAGTGGCTGCTGTAGTTGTTATGGCTTGCAATCGGGCTCATTAACCTTGGAAAAAGACT	498
Db	441	ATGCTCTGGCTGCTGTAGTGGTTATAGGCTGAGTCTGTCAGACTATCTTGAAGGACT	500
Qy	499	ATTAAATCCATCTTAAATATCCAAATATCTGTTGCCGTCAAATATCCTCTTTTCATATCC	558
Db	501	GTTAAATCAGTTTAACATATCAAACTCCCGTTGTTCAAATATCCTCTATTATATCT	560
Qy	559	CAGATGGATCANAATCTGATGTCAGAAAGCTTCTTTGAGCTATGATCAGCTGACGTAT	618
Db	561	CAGATGGATCTGATCAAGCTGTCAAGAGCAAGTCAATTGAGCTATTAATCAATTAACATAT	620
Qy	619	ATGCAGACATTTGAGATTTTGAACCTGTGCATCTGAAAGACACGGGAGCTGATGTCATAC	678
Db	621	ATGCAGACATTTGAGATTTTGAACCAAGTGGTCACTGAAAGGCTGGTGAACCTGCGTAC	680
Qy	679	TACAAAATTCGACGTCATTAAGTGGGCATTTGGATCAGCTGTTTTCACAAAGCAATAATTT	738
Db	681	TACAAGATTGCAGCTCACTCAAGTGGGCACCTGGACCAGTTGTTTTACAAACACAAATTT	740
Qy	739	AGCCGTGTTATCATCTAGAAAGATGATATGGAATATGCCCTGATTTTTTTTGGACTTTTT	798
Db	741	AGTCGAGTGATTATCTAGAAAGCATATGGAATTTGCTCCAGACTTCTTTGATTACTTT	800
Qy	799	GAGCTGGAGCTACTCTTCTTGACAGAGACAGTCGATATGCTATTTCTCTCTGGAAT	858
Db	801	GAGCTGCAGCTAGTCTCATGATAGGATAAACCAATTAATGCTGCTTCATCATGGAAT	860
Qy	859	GACAAATGGACAAATTCGAGTTTGTCCAAAGATCCTTATGCTCTTTTACCCTCAGATTTTTT	918
Db	861	GATTAATGGACAGACGATTTGTGCAATGCCCTATGGCTATACCGATCAGATTTTTT	920
Qy	919	CCCGCTCTTGGATGGATGCTTTCAAAATCTACTTTGGGACGAAATATCTCAAAAGTGGCCA	978
Db	921	CCTGGCTTTGGGTGGATGCTCAAGAGATCGACTTTGGGATGAGTTATCACAAAGTGGCCA	980
Qy	979	AAGCTTTACTGGGACGACTGGCTAAGACTCAAGAGAAATCACAGAGTGCACAAATTAAT	1038
Db	981	AAGGCTTACTGGGATGATTTGGCTCAGACTAAAGGAAACCAATAAGGCCGCAATTCAT	1040
Qy	1039	CGCCAGAAATTTGCAGAAACATAAATTTTGGTGAGCATGGTCTTAGTTTGGGGCAGTTT	1098
Db	1041	GCACGGAAAGTCTGTAGAACATAAATTTTGGTGAACATGGGTCTAGTTTGGGACAGTTT	1100
Qy	1099	TTCAAGCAGTATCTTGAGCCAAATTAACATAATGATGTCCAGGTTGATTTGGAAGTCAATG	1158
Db	1101	TTCACTCAGTATCTTGGAACTTATAAGCTAAACGATGTGACGGTTGACTGGAAGCAAG	1160
Qy	1159	GACCTTAGTTACTTTTGGAGGACAAATTAACGTGAAACACTTTTGGTGTGACTTGGTTAAAAG	1218
Db	1161	GACCTGGGATACCTGCACAGAGGGAACTATACCAAGTACTTTTCTGGCTTAGTGAGACAA	1220
Qy	1219	GCTAAGCCCATCTCAATGGAGCTGATGCTGCTTTGAAAGCATTTAAACATAGATGGTGAATG	1278
Db	1221	GCACGACCAATTCAGGTTCTGACCTTGTCTTAAAGGCTCAAAACATAAAGGATGATGAT	1280
Qy	1279	CGTATTCAGTACAGAGATCAACTAGACTTTGAAAATATCGCACCGGCAATTTTGGCAATTTT	1338

Db	1281	CGTATCCGGTATAAAGACCAAGTAGAGTTTGAACGCATTGCAGGGGAATTTGGTATATTT	1340
Qy	1339	GAAGAATGGAAGGATGGTGTACCACTGGCAGCATATAAAGGAATAGTAGTTTCCGGTAC	1398
Db	1341	GAAGAATGGAAGGATGGTGTGCCACCAACAGCATATAAAGGAGTAGTGGTGTTCGAATC	1400
Qy	1399	CAACGTCAGACGCTGTATTCCTTGTGTGGCCATGATTCGCTTCAACAACTCGGAATTGAA	1458
Db	1401	CAGACAAACAGACGCTGTATTCCTTGTGTGGCCAGATTCTGTAATGCAGCTTGGAAATTCGA	1460
Qy	1459	GATACTTTAACAAAGATATGATTCAGGA	1486
Db	1461	AATCTCGATGCAAAACATATGAAGGA	1488

RESULT 4

US-09-806-708B-22

; Sequence 22, Application US/09806708B

; Patent NO. 6784342

; GENERAL INFORMATION:

; APPLICANT: The University of British Columbia

; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants

; FILE REFERENCE: 4810-58741

; CURRENT APPLICATION NUMBER: US/09/806,708B

; CURRENT FILING DATE: 2001-04-03

; PRIOR APPLICATION NUMBER: US 60/147,133

; PRIOR FILING DATE: 1999-08-04

; NUMBER OF SEQ ID NOS: 23

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 22

; LENGTH: 1141

; TYPE: DNA

; ORGANISM: Artificial sequence

; FEATURE:

; NAME/KEY: promoter

; LOCATION: (1)..(1141)

; OTHER INFORMATION: Consensus sequence of A.t., L.a., and B.n. FAEL promoter

US-09-806-708B-22

5

```
Db 465 CTNNNTTWKMKTYNNCKWMSWNGSHRBAABAVTYWWMWRRYAHANNNNWDYWKKA 524
Qy 873 GCAGTTTGTCAAGATCCTTATGCTCTTTACCGCTCAGATTTTTCCTCCGGCTTGGATG 932
Db 525 CTWYKYBVCWKNNYAAVYTKSSWNTSYRYRKTNNSWRWSDTRSMGRANNYARABH 584
Qy 933 GATCGTTTCAAAATCTACTTGGGA--CGAATTATCTCCAAGTGGCGCAAGGCTTACTG 989
Db 585 YGYKWNTRWBWSHTWBHBRAGAAYWBMWYBAKCHMKAWYKAKKYAGAGGSNNNNNN 644
Qy 990 GGACGACTGGCTAGACTCAAGAGAGATCACAGAGGTCGACAAATTTATTCGCCCAGAGT 1049
Db 645 NNNNNNNNATCARDIYAAASRYAMANAkWYYKBAANNAYYTHANNWGCWNNATDT 704
Qy 1050 TTGCAGAACATATAATTTTGGTGAGCATGTTCTAGTTTGGGGCAGTTTTTCAAGCAGTA 1109
Db 705 RRTWKNNNNNNAGTWKNNNNNAKNAASAKNYAAAATKAAKCHRWANKWAMRGHADA 764
Qy 1110 TCTTGAG 1116
Db 765 AABTTDK 771
```

## RESULT 5

```
US-09-806-708B-22/c
; Sequence 22, Application US/09806708B
; Patent No. 6784342
; GENERAL INFORMATION:
; APPLICANT: The University of British Columbia
; TITLE OF INVENTION: Regulation of Embryonic Transcription in Plants
; FILE REFERENCE: 4810-58741
; CURRENT APPLICATION NUMBER: US/09/806,708B
; PRIOR FILING DATE: 2001-04-03
; PRIOR FILING DATE: US 60/147,133
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22
; LENGTH: 1141
; TYPE: DNA
; ORGANISM: Artificial sequence
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (1)..(1141)
; OTHER INFORMATION: consensus sequence of A.t., L.a., and B.n. FAEL promoters
US-09-806-708B-22
```

```
Query Match 2.6%; Score 45.4; DB 3; Length 1141;
Best Local Similarity 10.8%; Pred. No. 0.0018;
Matches 63; Conservative 202; Mismatches 317; Indels 0; Gaps 0;

Qy 764 ATATGGAAATGCCCTCGATTTTTCGACTTTTGTGAGGCTGGAGTACTCTTCTTGACA 823
Db 915 WTACYNRAATNNKMATWMMKTHGASHKBRTHRTCRRTKYNNNNNNARTVYVYHAA 856
Qy 824 GAGACAAGTCAGTATGCTATTTCTTCTTGGAAATGACAATGGACAAATGACGTTTGTCC 883
Db 855 RRRWNAWTRTNNNNNNNNNACNRTWTWABWKHSWCNNNNNNNNNNNNNNNNNTWCHYTTA 796
Qy 884 AAGATCCTTATGCTCTTTACCGCTCAGATTTTTCCTCCGGTCTTGGATGGATGCTTCAA 943
Db 795 NABECYRANNNAARWARTCNNTWHAATVTTHTDWCYKTMWNTWYDMMTMBTTTTRN 736
Qy 944 AATCTACTTGGGACGAATTAATCTCCAAAGTGGCCAAAGGCTTACTGGGACGACTGGCTAA 1003
Db 735 MTSTNNNNNNWACTNNNNNNMKAYAHATNKGWGNWNTDARRNNNTVMERRW 676
Qy 1004 GACTCAAGAGAAATCACAGAGGTCGACAAATTTATTCGCCCGAAGATTTGCAGAACATATA 1063
Db 675 MTNTKTRWYSTTRRHVYTGATNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 616
Qy 1064 ATTTTGGTGAGCATGGTCTTAGTTTGGGGCAGTTTTTCAAGCAGCATATCTTGGACCAATTA 1123
```

```
Db 615 KKVKMRDITTCYVDVWADSWWYANWMECRDVTYTRNNYCKSYAHSYWWSNNAMWYR 556
Qy 1124 AACTAAATGATGTCAGGTTGATTTGGAAGTCAATAGGACCTTAGTTACCTTTTGGAGACA 1183
Db 555 RVSARNWSSMARWTTTRNNWMSGBVRMRWAGTMMWRHNNNNNTDTRYWMMKWARBTMT 496
Qy 1184 ATTACGTGAACACATTTTGGTGACTTGGTTTAAAGAGCTAAGCCCATCCATGGAGCTGATG 1243
Db 495 VYDSMCNAKSMWRGNWRAKMMWMAANNNDAGAMDHWTYWGNNTMMRRRAKMMNMAWCR 436
Qy 1244 CTGCTCTTGAAGACATTTAAACATAGATGCTGATTCGCGTATTCAGTAGCAGAGATCAACTAG 1303
Db 435 RAVCCNNNNNRACVWHKHKMRWTWYMKKACACNNNNBRKAMYRVAMWYSRDNTTNDWM 376
Qy 1304 ACTTTGAAATATCGCAGCGCAATTTGGCAATTTTGAAGAAT 1345
Db 375 MMTSDWBWHWYTVDYTMRAWNNNNNNNNNNNRBCKTTSWMMWMD 334
```

## RESULT 6

```
US-09-270-767-9920/c
; Sequence 9920, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9920
; LENGTH: 778
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-9920
```

```
Query Match 2.3%; Score 40; DB 3; Length 778;
Best Local Similarity 47.6%; Pred. No. 0.075;
Matches 118; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

Qy 1155 AATGGACCTTAGTTACCTTTTGGAGGACAATACGTTGGAACACTTTGGTGACTTGGTTAA 1214
Db 756 AATCAATATGAGTACTCTGCTGAAGACAATACGATAACACAGTTTCTTGGCGGCGGTTTA 697
Qy 1215 AAAGGCTAAGCCCATCCATGGAGCTGATGCTGTTTGAAGAGCATTTAAACATAGATGTGA 1274
Db 696 TAGCTATCCATTTGTTACGTACGATGAGCTGCGCGGAACCTGATTAGATCGAAGTCC 637
Qy 1275 TGTGCGTATTCAGTACAGAGATCAACTAGACTTTTGAATAATATCGACGGCAATTTGGCAT 1334
Db 636 AGTTCGATTCAATACACTACTAGGAGCAGTACAAGCGGACAACTAAGATGCTGGGTCT 577
Qy 1335 TTTTGAAGATGAAGGATGGTGACCACTGACGTCAGCATATAAGGAATAGTAGTTTCCG 1394
Db 576 TATGGATGATTTCAAGAGCGGTGTTCCCGGACTGCTACCATGGCATGCTCTCTTCTA 517
Qy 1395 GTACCAAA 1402
Db 516 CTACAACA 509
```

## RESULT 7

```
US-09-134-001C-2477
; Sequence 2477, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
```

```
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 2477
; LENGTH: 2793
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
; US-09-134-001C-2477

Query Match      2.2%; Score 38.8; DB 3; Length 2793;
Best Local Similarity 51.8%; Pred. No. 0.41;
Matches 88; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

Qy 1213 AAAAAGGCTAAGCCCATCCATGGAGCTGATGCTCTCTTGAAGCAATTTAAACATAGATGGT 1272
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2035 ACAATCTCGCACTAGACTCAGAAGGCAATTCCTGATGGGTAACTATTAAACGGACAAGAC 2094
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1273 GATGTGGGTATTAGTACAGAGATCAACTAGACTTTGAAATATCGCACGCAATTTGGC 1332
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2095 GATATGACTATTGATAGCGGATTTTATCAACACACTAAATATAGCTTAGGGAATATGTA 2154
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1333 ATTTTGAAGATGGAAGGATGGTGATCCACGTCGACGATATAAAGGAAT 1382
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2155 TGGTATGACACTAATAAGATGGTATTCAAGGTGATGATGAAGGAAT 2204
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 8
US-09-386-962C-7
; Sequence 7, Application US/09386962C
; Patent No. 6635473
; GENERAL INFORMATION:
; APPLICANT: POSTER, Timothy
; TITLE OF INVENTION: POLYPEPTIDES AND POLYNUCLEOTIDES FROM COAGULASE-NEGATIVE STAPHYLOCOCCUS
; FILE REFERENCE: P06335US2/BAS
; CURRENT APPLICATION NUMBER: US/09/386,962C
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: 60/098,443
; PRIOR FILING DATE: 1998-08-31
; PRIOR APPLICATION NUMBER: 60/117,119
; PRIOR FILING DATE: 1999-01-25
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 7
; LENGTH: 2976
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (3)..(2975)
; US-09-386-962C-7

Query Match      2.2%; Score 38.8; DB 3; Length 2976;
Best Local Similarity 51.8%; Pred. No. 0.41;
Matches 88; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

Qy 1213 AAAAAGGCTAAGCCCATCCATGGAGCTGATGCTCTTGAAGCAATTTAAACATAGATGGT 1272
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2136 ACAATCTCGCACTAGACTCAGAAGGCAATTCCTGATGGGTAACTATTAAACGGACAAGAC 2195
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1273 GATGTGGGTATTAGTACAGAGATCAACTAGACTTTGAAATATCGCACGCAATTTGGC 1332
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2196 GATATGACTATTGATAGCGGATTTTATCAACACACTAAATATAGCTTAGGGAATATGTA 2255
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1333 ATTTTGAAGATGGAAGGATGGTGATCCACGTCGACGATATAAAGGAAT 1382
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2256 TGGTATGACACTAATAAGATGGTATTCAAGGTGATGATGAAGGAAT 2305
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 9
US-09-386-959-7
; Sequence 7, Application US/09386959
; Patent No. 6703025
; GENERAL INFORMATION:
; APPLICANT: PATTI, Joseph M.
; APPLICANT: FOSTER, Timothy J.
; APPLICANT: HOOK, Magnus
; TITLE OF INVENTION: MULTICOMPONENT VACCINES
; FILE REFERENCE: P06333US1/BAS
; CURRENT APPLICATION NUMBER: US/09/386,959
; CURRENT FILING DATE: 1999-08-31
; EARLIER APPLICATION NUMBER: 60/098,439
; EARLIER FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 7
; LENGTH: 2976
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (4)..(2976)
; US-09-386-959-7

Query Match      2.2%; Score 38.8; DB 3; Length 2976;
Best Local Similarity 51.8%; Pred. No. 0.41;
Matches 88; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

Qy 1213 AAAAAGGCTAAGCCCATCCATGGAGCTGATGCTCTTGAAGCAATTTAAACATAGATGGT 1272
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2136 ACAATCTCGCACTAGACTCAGAAGGCAATTCCTGATGGGTAACTATTAAACGGACAAGAC 2195
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1273 GATGTGGGTATTAGTACAGAGATCAACTAGACTTTGAAATATCGCACGCAATTTGGC 1332
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2196 GATATGACTATTGATAGCGGATTTTATCAACACACTAAATATAGCTTAGGGAATATGTA 2255
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1333 ATTTTGAAGATGGAAGGATGGTGATCCACGTCGACGATATAAAGGAAT 1382
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2256 TGGTATGACACTAATAAGATGGTATTCAAGGTGATGATGAAGGAAT 2305
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
; Sequence 7, Application US/09386959
; Patent No. 6703025
; GENERAL INFORMATION:
; APPLICANT: PATTI, Joseph M.
; APPLICANT: FOSTER, Timothy J.
; APPLICANT: HOOK, Magnus
; TITLE OF INVENTION: MULTICOMPONENT VACCINES
; FILE REFERENCE: P06333US1/BAS
; CURRENT APPLICATION NUMBER: US/09/386,959
; CURRENT FILING DATE: 1999-08-31
; EARLIER APPLICATION NUMBER: 60/098,439
; EARLIER FILING DATE: 1999-08-31
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 7
; LENGTH: 2976
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (4)..(2976)
; US-09-386-959-7

Query Match      2.2%; Score 38.8; DB 3; Length 2976;
Best Local Similarity 51.8%; Pred. No. 0.41;
Matches 88; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

Qy 1213 AAAAAGGCTAAGCCCATCCATGGAGCTGATGCTCTTGAAGCAATTTAAACATAGATGGT 1272
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2136 ACAATCTCGCACTAGACTCAGAAGGCAATTCCTGATGGGTAACTATTAAACGGACAAGAC 2195
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1273 GATGTGGGTATTAGTACAGAGATCAACTAGACTTTGAAATATCGCACGCAATTTGGC 1332
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2196 GATATGACTATTGATAGCGGATTTTATCAACACACTAAATATAGCTTAGGGAATATGTA 2255
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 1333 ATTTTGAAGATGGAAGGATGGTGATCCACGTCGACGATATAAAGGAAT 1382
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 2256 TGGTATGACACTAATAAGATGGTATTCAAGGTGATGATGAAGGAAT 2305
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 10
US-09-248-796A-2786/c
; Sequence 2786, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 2786
; LENGTH: 942
; TYPE: DNA
; ORGANISM: Candida albicans
; US-09-248-796A-2786

Query Match      2.2%; Score 38.2; DB 3; Length 942;
Best Local Similarity 54.7%; Pred. No. 0.31;
Matches 76; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

Qy 751 ATACTAGAGATGATATGGAATTTGCCCTCGATTTTTTTTGACTTTTTTTGAGCTGAGCT 810
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 196 ATTCAAATGTTGATATCGTTTGGCCATGATTTTCCTGATGTGTTGACTGTGGGGT 137
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 811 ACTCTTCTTGACAGACAAAGTCGATTTATGGCTATTTCTTCTTGAATGACATGACNA 870
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 136 GTTCTCTTGATCAATAAAGAAATTGTATAGGAGATTTATTATTATAAAGCGATGATCAT 77
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

QY 871 ATGCAGTTTGCCAGATC 889  
Db 76 AATCTTATGAATAAATC 58

## RESULT 11

US-09-949-016-16440/C  
; Sequence 16440, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16440  
; LENGTH: 32172  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)..(32172)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-16440

Query Match 2.1%; Score 36.8; DB 3; Length 32172;  
Best Local Similarity 48.1%; Pred. No. 7.8;  
Matches 104; Conservative 0; Mismatches 112; Indels 0; Gaps 0;  
QY 301 CAGATTAGCTGCAGCAAGAAATAGTTGCTCTTGAAGAACAAATGAAGCGTCAGGAC 360  
Db 18939 CAGATCATTTTGACTCTTGTGCCTAGAAATGTACTTGGCACATAGCAGAAACTCAAGCAC 18880  
QY 361 CAGGAGTCCGACAAATTAAGGCTCTTGTTCAGACTCTTGAAGTAAGGCATAAAAAG 420  
Db 18879 ATTATTTAGTGAATAAATTTGCTGCTTTCTATAGTAATAAACAATTCATAAGAAATG 18820  
QY 421 TTGATCGAAATGTACAGATGCCAGTGGCTGCTAGTTGTTATGCTTGCATTCGGGCT 480  
Db 18819 TTTAACTTAATGTATTAACGTCTCTAACAGCTTATGTGTTATGTGGATTAGTAGGCA 18760  
QY 481 GATTACCTGGAAGAGACTATTAAATCCATCTTAAAA 516  
Db 18759 GGTATATTGAATAGGTTAATTTATCAGCTAGAAAA 18724

## RESULT 12

US-09-620-312D-194  
; Sequence 194, Application US/09620312D  
; Patent No. 6569662  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Zhang, Jie  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Chen, Rui-hong  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Wehrman, Tom  
; APPLICANT: Xue, Aidong J.  
; APPLICANT: Yang, Yonghong  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Zhou, Ping

; APPLICANT: Ma, Yunging  
; APPLICANT: Wang, Dunrui  
; APPLICANT: Wang, Zhiwei  
; APPLICANT: John Tillinghaast  
; APPLICANT: Drihanac, Radoje T.  
; TITLE OF INVENTION: Polypeptides  
; FILE REFERENCE: 784CIP2B  
; CURRENT APPLICATION NUMBER: US/09/620,312D  
; CURRENT FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: 09/552,317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 1105  
; SOFTWARE: pt FL\_genes Version 1.0  
; SEQ ID NO 194  
; LENGTH: 2353  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (153)..(2309)  
US-09-620-312D-194

Query Match 2.1%; Score 36.6; DB 3; Length 2353;  
Best Local Similarity 49.2%; Pred. No. 1.8;  
Matches 96; Conservative 0; Mismatches 99; Indels 0; Gaps 0;  
QY 744 TGTATCATACTAGAAAGATGATGAAATGGCCCTGATTTTTTGTGAGGC 803  
Db 1316 TGCTGGTCTTGGAAAGAGACCTGGACATTTGCTGGATTTTTCAGTTTCTCTGAGCCA 1375  
QY 804 TGGAGCTACTCTTTTGACAGAGACAAGTCGATTTATGGCTATTCTTTTGAATGACAA 863  
Db 1376 ATCCATCCACCTACTGGAGGAGATGACAGCCCTGTACTGTCATCTCTGCTTGAATGACCA 1435  
QY 864 TGGACAAATGCAGTTTGTCCAGATCCTTATGCTCTTTACCGCTCAGATTTTTTCCCGG 923  
Db 1436 GGGGTATGAACACACACCGCTGAGGACCCAGCACTACTGTACCGTGTGGAGACCATGCGCTGG 1495  
QY 924 TCTTGATGGATGCT 938  
Db 1496 CTGGGCTGGTGCT 1510

## RESULT 13

US-10-012-231A-87  
; Sequence 87, Application US/10012231A  
; Patent No. 6924355  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Botstein, David  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Eaton, Dan I.  
; APPLICANT: Ferrara, Napoleone  
; APPLICANT: Fong, Sherman  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Grimaldi, Christopher J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Hillan, Kenneth J.  
; APPLICANT: Pan, James  
; APPLICANT: Paoni, Nicholas F.  
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic  
; Acids Encoding the Same  
; FILE REFERENCE: P2830PIC23  
; CURRENT APPLICATION NUMBER: US/10/012,231A  
; CURRENT FILING DATE: 2002-06-10  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 477  
; SEQ ID NO 87

Search completed: December 14, 2005, 13:08:39  
Job time : 343 secs

This Page Blank (uspio)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2005, 06:14:15 ; Search time 1005 Seconds

(without alignments)  
11518.976 Million cell updates/sec

Title: US-10-615-144-3

Perfect score: 1737

Sequence: 1 gaattcgccgcattga.....ataatcgccgcgaattc 1737

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq 21.\*

- 1: geneseqn1980s.\*
- 2: geneseqn1990s.\*
- 3: geneseqn2000s.\*
- 4: geneseqn2001as.\*
- 5: geneseqn2001bs.\*
- 6: geneseqn2002as.\*
- 7: geneseqn2002bs.\*
- 8: geneseqn2003as.\*
- 9: geneseqn2003bs.\*
- 10: geneseqn2003cs.\*
- 11: geneseqn2003ds.\*
- 12: geneseqn2004as.\*
- 13: geneseqn2004bs.\*
- 14: geneseqn2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1737	100.0	1737	2	Aax78002 Tobacco G
2	1255	72.3	1669	2	Aax78001 Potato Gn
3	770.4	44.4	1639	3	Aac39540 Arabidops
4	748	43.1	1641	3	Aac50197 Arabidops
5	714.2	41.1	1836	13	Adx54592 Plant ful
6	661.8	38.1	1882	13	Ado81767 Plant ful
7	661	38.1	1211	13	Adr61718 Cotton cd
8	658.6	37.9	2180	10	Add68117 Glycosylt
9	658.6	37.9	9240	10	Add68117 Glycosylt
10	658	37.9	1319	10	Add68120 Glycosylt
11	656.6	37.8	1119	10	Add68136 Plant Gnt
12	581.2	33.5	1898	13	Adt20185 Plant cDN
13	425.8	24.5	896	14	Aeb55577 Rice geno
14	334	19.2	510	2	Aax78003 A. thalia
15	182	10.5	278	7	AdS70553 Corn seed
16	181.8	10.5	2580	11	Acn44097 Mouse mRN
17	181.8	10.5	31477	11	Acn44096 Mouse gen
18	181.6	10.5	2546	10	AdB58546 Toxicity-
19	181.6	10.5	2546	10	AdB53154 Primary r

20	181.6	10.5	2557	2	AAQ62625	Aaq62625 Rat N-ace
21	165.6	9.5	2485	2	AAQ25594	Aaq25594 Rabbit Gn
22	146.4	8.4	1335	4	AAI69837	Aai69837 Human bet
23	146.4	8.4	2518	13	ADT07465	Adt07465 Human col
24	146.4	8.4	2602	6	ABK84032	Abk84032 Human cDN
25	146.4	8.4	2614	13	ADT07459	Adt07459 Human col
26	146.4	8.4	2631	13	ADT07458	Adt07458 Human col
27	146.4	8.4	2670	12	ADJ57449	Adj57449 Human Glc
28	146.4	8.4	2670	14	ADZ69846	Adz69846 Human Gnt
29	146.4	8.4	2680	13	ADT07460	Adt07460 Human col
30	146.4	8.4	2688	10	ADI62841	Adi62841 Human apo
31	146.4	8.4	2688	10	ADI62862	Adi62862 Human apo
32	146.4	8.4	2753	11	ACN44099	Acn44099 Human mRN
33	146.4	8.4	2842	13	ADT07454	Adt07454 Human col
34	146.4	8.4	2900	13	ADT07464	Adt07464 Human col
35	146.4	8.4	2936	13	ADT07457	Adt07457 Human col
36	146.4	8.4	2937	13	ADT07430	Adt07430 Human Gnt
37	146.4	8.4	2937	14	ADY14405	Ady14405 DNA encod
38	146.4	8.4	2937	14	ADZ49000	Adz49000 Insulin s
39	146.4	8.4	3067	13	ADT07463	Adt07463 Human col
40	146.4	8.4	3106	13	ADT07462	Adt07462 Human col
41	146.4	8.4	3135	13	ADT07455	Adt07455 Human col
42	146.4	8.4	3183	13	ADT07456	Adt07456 Human col
43	146.4	8.4	3230	2	AAQ25595	Aaq25595 Human Gnt
44	146.4	8.4	4677	12	ADJ57451	Adj57451 Vector pp
45	146.4	8.4	4677	14	ADZ69848	Adz69848 Plasmid p

ALIGNMENTS

RESULT 1  
AAX78002  
ID AAX78002 standard; cDNA; 1737 BP.

AC AAX78002;

DT 19-AUG-1999 (first entry)

DE Tobacco GntI cDNA.

XX GntI; beta-1,2-N-acetylglucosaminyltransferase; glycoprotein; plant;  
KW deficient; defective; detection; transgenic plant; sugar residue;  
KW medicine; antigenicity; deglycosylation; tobacco; ds.  
XX Nicotiana tabacum.

XX Key Location/Qualifiers  
FH 127..1467  
CDS /\*tag= a

FT /product= "GntI"

XX DE19754622-A1.

XX PD 10-JUN-1999.

XX PF 09-DEC-1997; 97DE-01054622.

XX PR 09-DEC-1997; 97DE-01054622.

XX PA (VSCH/) VON SCHAEWEN A.

XX PI Von Schawen A;

XX DR WPI; 1999-338905/29.

XX PT P-PSDB; AAY08889.

XX Nucleic acid encoding plant N-acetylglucosaminyl transferase I useful for  
generating plants producing glycosylated proteins of low antigenicity.

XX Claim 7; Page 22-24; 37pp; German.

XX This invention describes novel plant N-acetylglucosaminyl transferase I

= priority doc  
in instant  
appl.

CC (GnTI) proteins and their encoding nucleic acids. The nucleic acid  
CC sequences of the invention may be used for recombinant production of the  
CC encoded proteins, which are then used to raise antibodies (Ab) for  
CC identifying plants with deficient or defective production of the protein.  
CC They may also be used to detect such plants by hybridization and to  
CC isolate related sequences from other plants or to generate antisense or  
CC sense constructs for reducing/deleting GnTI protein activity in plants.  
CC These transgenic plants may be used to produce glycoproteins with  
CC minimal, uniform and defined sugar residues. Such glycoproteins are  
CC useful in medicine and research, e.g. human glucocerebrosidase for  
CC treating Gaucher's disease. Plants which are defective or deficient in  
CC production of the GnTI protein can be made to produce glycoproteins with  
CC minimal, uniform and defined sugar residues, of low antigenicity. Use of  
CC these plants eliminates the need for the difficult isolation and  
CC deglycosylation of native proteins or preparation in defective animal  
CC cells. This sequence represents tobacco (Nicotiana tabacum) GnTI encoding  
CC cDNA  
XX  
SQ Sequence 1737 BP; 510 A; 337 C; 390 G; 500 T; 0 U; 0 Other;  
  
Query Match 100.0%; Score 1737; DB 2; Length 1737;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1737; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 GAATTCGGCGCGCATTTGACTGCTTAACCTGAACAGGCAAGTAATCCAGCGATGA 60  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 1 GAATTCGGCGCGCATTTGACTGCTTAACCTGAACAGGCAAGTAATCCAGCGATGA 60  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 61 AACACTCAATCACTGAACACTGAGAGACTATTGCGTTCTCTTAAGGCTTCAATCGAATT 120  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 61 AACACTCAATCACTGAACACTGAGAGACTATTGCGTTCTCTTAAGGCTTCAATCGAATT 120  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 121 CGCAGATGAGAGGAAACAAGTTTGTGTGATTTCCGGTACCTCTCATCTTGGCTGCT 180  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 121 CGCAGATGAGAGGAAACAAGTTTGTGTGATTTCCGGTACCTCTCATCTTGGCTGCT 180  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 181 GTCGCTTCACTACACAGATCGGCTTTTTCGGACACAGTCAGATATGCAATCGCATCGC 240  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 181 GTCGCTTCACTACACAGATCGGCTTTTTCGGACACAGTCAGATATGCAATCGCATCGC 240  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 241 CTTGCTGCTGCAATTGAAGCAGAAATCATTTGTACAGGACGACAGATGCTTATTGAC 300  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 241 CTTGCTGCTGCAATTGAAGCAGAAATCATTTGTACAGGACGACAGATGCTTATTGAC 300  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 301 CAGATTAGCTGCGACGAGGAAGTAATAGTTGCTCTTGAAGAACAAATGAAGCGTCAGGAC 360  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 301 CAGATTAGCTGCGACGAGGAAGTAATAGTTGCTCTTGAAGAACAAATGAAGCGTCAGGAC 360  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 361 CAGAGTGCCGACAAATTAAGGGCTCTTTGTCAGGATCTTGAAGTAAGGGCATAAAG 420  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 361 CAGAGTGCCGACAAATTAAGGGCTCTTTGTCAGGATCTTGAAGTAAGGGCATAAAG 420  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 421 TTGATCGGAATGTACAGATGCCAGTGGCTGTCTAGTTGTTATGGCTTGCATTCGGGCT 480  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 421 TTGATCGGAATGTACAGATGCCAGTGGCTGTCTAGTTGTTATGGCTTGCATTCGGGCT 480  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 481 GATTAACCTGGAAAAAGACTATTAATCCATCTTAAATAACCAATATCTGTTCGGTCAAAA 540  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 481 GATTAACCTGGAAAAAGACTATTAATCCATCTTAAATAACCAATATCTGTTCGGTCAAAA 540  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 541 TATCCTCTTTTCATATCCAGGATGGATCATCTCATCTGATGTCAGGAGCTTGTTCGAGC 600  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 541 TATCCTCTTTTCATATCCAGGATGGATCATCTCATCTGATGTCAGGAGCTTGTTCGAGC 600  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 601 TATGATCAGCTGATATGACAGCACTTGGATTTTGAACCTGTGTCATCTGAAAGACCA 660  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 601 TATGATCAGCTGATATGACAGCACTTGGATTTTGAACCTGTGTCATCTGAAAGACCA 660  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 661 GGGAGCTGATGTCATCTACAAAATTGCAAGTTCAGCTCATTTACAGTGGGCAATTCAGCTG 720  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 661 GGGAGCTGATGTCATCTACAAAATTGCAAGTTCAGCTCATTTACAGTGGGCAATTCAGCTG 720  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
QY 721 TTTTACAAGCATAATTTTAGCCGTGTTATCATACTAGNAGATGATAGGAATTTGCCCT 780

RESULT 2  
AAK78001  
ID AAK78001 standard; cDNA; 1669 BP.  
XX  
AC AAK78001;

Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 780  
QY TTTTACAAGCATAATTTTAGCCGTGTTATCATCTAGAAGATGATAGGAATTTGCCCT 780  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 840  
QY 781 GATTTTTTTGACTTTTTTTGAGGCTGAGCTACTCTTTCTTGACAGACAAAGTCGATTATG 840  
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 840  
Db 781 GATTTTTTTGACTTTTTTTGAGGCTGAGCTACTCTTTCTTGACAGACAAAGTCGATTATG 840  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 900  
QY 841 GCTATTTCTCTTGGAAATGCAATGCAAAATGCAAGTTTGTCCAAAGATCCTTATGCTCTT 900  
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 900  
Db 841 GCTATTTCTCTTGGAAATGCAATGCAAAATGCAAGTTTGTCCAAAGATCCTTATGCTCTT 900  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 960  
QY 901 TACCGCTCAGATTTTTTTCCCGGCTCTTGATGATGATCTTTCAAAATCTACTTGGGACGAA 960  
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 960  
Db 901 TACCGCTCAGATTTTTTTCCCGGCTCTTGATGATGATCTTTCAAAATCTACTTGGGACGAA 960  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 1020  
QY 961 TTATCTCCAAAGTGGCCAAAGGCTTACTCGGACGACTGCTAAGACTCAAGAGAAATCAC 1020  
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 1020  
Db 961 TTATCTCCAAAGTGGCCAAAGGCTTACTCGGACGACTGCTAAGACTCAAGAGAAATCAC 1020  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 1080  
QY 1021 AGAGTTCGACAAATTTATTCGCCGAGAGTTTGCAGAACATATAATTTTGGTGAAGCATGTT 1080  
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 1080  
Db 1021 AGAGTTCGACAAATTTATTCGCCGAGAGTTTGCAGAACATATAATTTTGGTGAAGCATGTT 1080  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 1140  
QY 1081 TCTAGTTTTCGGGCGAGTTTTTCAAGCAGATATCTTGAGCCAAATTAACCTAAATGATGTCAG 1140  
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 1140  
Db 1081 TCTAGTTTTCGGGCGAGTTTTTCAAGCAGATATCTTGAGCCAAATTAACCTAAATGATGTCAG 1140  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 1200  
QY 1141 GTTGATTGGAAGTCAATGGACCTTAGTTACTCTTTTGGAGGACAAATTTACGTGAACACTTT 1200  
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 1200  
Db 1141 GTTGATTGGAAGTCAATGGACCTTAGTTACTCTTTTGGAGGACAAATTTACGTGAACACTTT 1200  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 1260  
QY 1201 GGTGACTTGGTTTAAAGGCTAAGCCCATCCATGGAGCTGATGCTGCTTTGAAAGCATTT 1260  
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 1260  
Db 1201 GGTGACTTGGTTTAAAGGCTAAGCCCATCCATGGAGCTGATGCTGCTTTGAAAGCATTT 1260  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 1320  
QY 1261 AACATAGATGTTGATGTTGCGTATTCAGTACAGAGATCAACTGACATTTTGAAATATCGCA 1320  
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 1320  
Db 1261 AACATAGATGTTGATGTTGCGTATTCAGTACAGAGATCAACTGACATTTTGAAATATCGCA 1320  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 1380  
QY 1321 CGGCAATTTGGCAATTTTGAAGAAATGGAAGGATGGTACCACGTGCAGCATATAAAGGA 1380  
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 1380  
Db 1321 CGGCAATTTGGCAATTTTGAAGAAATGGAAGGATGGTACCACGTGCAGCATATAAAGGA 1380  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 1440  
QY 1381 ATAGTATGTTTTCGGTACAAAGCTCCAGCGTGTATTCCTTGTGGCCCATGATTCGCTT 1440  
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 1440  
Db 1381 ATAGTATGTTTTCGGTACAAAGCTCCAGCGTGTATTCCTTGTGGCCCATGATTCGCTT 1440  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 1500  
QY 1441 CAACAACTCGGAATTCGAAGATCTTAACAAAGATATGTTTGCAGGAGCCCGGGCAAAATTT 1500  
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 1500  
Db 1441 CAACAACTCGGAATTCGAAGATCTTAACAAAGATATGTTTGCAGGAGCCCGGGCAAAATTT 1500  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 1560  
QY 1501 TTTGACTTATTTGGGTAGGATGTCATCGAGCTGCACCTAAACCATGATTTTACCAAGTTACAT 1560  
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 1560  
Db 1501 TTTGACTTATTTGGGTAGGATGTCATCGAGCTGCACCTAAACCATGATTTTACCAAGTTACAT 1560  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 1620  
QY 1561 ACAAAGTTTATGTTTATAGGAGGAGCTCACTGTTCTAGTGTGTTGAAGGATATCGGCTT 1620  
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 1620  
Db 1561 ACAAAGTTTATGTTTATAGGAGGAGCTCACTGTTCTAGTGTGTTGAAGGATATCGGCTT 1620  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 1680  
QY 1621 CTTAGTATTTGGATGAAATCAACCAACCTATTATTTTAAAGTGTTCAGAACATAAAGAG 1680  
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 1680  
Db 1621 CTTAGTATTTGGATGAAATCAACCAACCTATTATTTTAAAGTGTTCAGAACATAAAGAG 1680  
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 1737  
QY 1681 GAAATGTAGCCCTGTAAAGACTATACATGGGACCATCATTAATCGCGCCCGCAATTC 1737  
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||| 1737  
Db 1681 GAAATGTAGCCCTGTAAAGACTATACATGGGACCATCATTAATCGCGCCCGCAATTC 1737



XX	19-AUG-1999	(first entry)	QY	306	TAGCCTGCAGCAAGGAAGAAATAGTCTCTTGAAGACAAATGAAGCGTCAGGACCAAGGA	365
DT			DB	232	TAGCCACGACGAGGAAGAGTAGTCTCTTGAAGACAAATGAAGCATCAGGACCAAGGA	291
DE	Potato GntI cDNA.					
XX	GntI; beta-1,2-N-acetylglucosaminyltransferase; glycoprotein; plant;		QY	366	GTGCGCAATTAAGGGCTCTTTCAGGATCTTGAAGTAAGGGCAATAAAGAGTTGAT	425
KW	deficient; detection; transgenic plant; sugar residue;		DB	292	GTGCGCGCAATTAAGGGCTCTTTCAGGATCTTGAAGTAAGGGCAATAAAGAGTTAAT	351
KW	medicine; antigenicity; deglycosylation; potato; ds.		QY	426	CGGAAATGTACAGATGCCAGTGGCTGTAGTTGTTATGGCTTGCATCGGGTGATTA	485
XX	Solanum tuberosum.		DB	352	CGGAGATGTGCAGATGCCAGTGGCAGCTGTAGTTGTTATGGCTTGCAGTCTGACTACTA	411
XX						
FH	Key	Location/Qualifiers	QY	486	CCTGGAAGAGACTATTAATCCATCTTAAATACCAATATCTCTTGGGTCAAAATATCC	545
FT	CDS	53..1394	DB	412	CCTGAGAGGACTATTAATCCATCTTAAATACCAATATCTCTTGGGTCAAAATATCC	471
FT		/*tag= a				
FT		/product= "GntI"				
XX	DE19754622-A1.		QY	546	TCCTTTTCATATCCAGGATGCATCATCTCTGATGTCAGGAAGCTTCTCTTGGCTATGA	605
PN			DB	472	TCCTTTTCATATCCAGGATGCATCATCTCTGATGTCAGGAAGCTTCTCTTGGCTATGA	531
PD	10-JUN-1999.					
XX	09-DEC-1997;	97DE-01054622.	QY	606	TCAGCTGACGTATATGACGACCTTGGATTTTGAACCTGTGTGCATCTCTGAAACAGCAGGGA	665
PF			DB	532	TCAGCTGACGTATATGACGACCTTGGATTTTGAACCTGTGTGCATCTCTGAAACAGCAGGGA	591
XX	09-DEC-1997;	97DE-01054622.	QY	666	GCTGATTGCATCTACAAAATGTCAGTCAATCAAGTGGGCATTTGGATGAGCTGTTTA	725
PA	(VSCH/) VON SCHAEWEN A.		DB	592	ACTGTTGCATCTACAAAATGTCAGTCAATCAAGTGGGCATTTGGATGAGCTGTTTA	651
XX	Von Schaewen A;					
XX	WPI; 1999-338905/29.		QY	726	CAAGCATTAATTTAGCCGTGTATCATCTACTAGAGATGATATGAAATTTGCCCTGATTT	785
DR	P-PSDB; AAY08888.		DB	652	CAAGCATTAATTTAGCCGTGTATCATCTACTAGAGATGATATGAAATTTGGCTGATTT	711
XX						
PT	Nucleic acid encoding plant N-acetylglucosaminyl transferase I useful for		QY	786	TTTTGACTTTTTGAGGCTGAGCTACTCTCTTTCGACAGACAAAGTGCATATGAGCTAT	845
PT	generating plants producing glycosylated proteins of low antigenicity.		DB	712	TTTTGACTTTTTGAGGCTGAGCTACTCTCTTTCGACAGACAAAGTGCATATGAGCTAT	771
XX	Claim 6; Page 16-18; 37pp; German.					
XX	This invention describes novel plant N-acetylglucosaminyl transferase I		QY	846	TTCTTTTGGAAATGACAAATGACAAATGCGATTTGTCTCAAGATCTTATGCTCTTTACCG	905
CC	(GntI) proteins and their encoding nucleic acids. The nucleic acid		DB	772	TTCTTTTGGAAATGACAAATGACAAATGCGATTTGTCTCAAGATCTTATGCTCTTTACCG	831
CC	sequences of the invention may be used for recombinant production of the					
CC	encoded proteins, which are then used to raise antibodies (Ab) for		QY	906	CTCAGATTTTTTCCCGCTCTTGGATGATCTCTTCAAAATCTCTTGGGACGATATTC	965
CC	identifying plants with deficient or defective production of the protein.		DB	832	CTCAGATTTTTTCCCGCTCTTGGATGATCTCTTCAAAATCTCTTGGGACGATATTC	891
CC	They may also be used to detect such plants by hybridization and to					
CC	isolate related sequences from other plants or to generate antisense or		QY	966	TCCAAAGTGGCCAAAGGCTTACTGGGACGCTGCTGCTAAGACTCAAAAGAGAAATCACAGAG	1025
CC	sense constructs for reducing/deleting GntI protein activity in plants.		DB	892	TCCAAAGTGGCCAAAGGCTTACTGGGACGCTGCTGCTAAGACTCAAAAGAGAAATCACAGAG	951
CC	These transgenic plants may be used to produce glycoproteins with		QY	1026	TCGCAATTTATTCGCCCCAGAGTTTGCAGAACATATTAATTTGGTGAGCATGGTTCTAG	1085
CC	minimal, uniform and defined sugar residues. Such glycoproteins are		DB	952	TCGCAATTTATTCGCCCCAGAGTTTGCAGAACATATTAATTTGGTGAGCATGGTTCTAG	1011
CC	useful in medicine and research, e.g. human glucocerebrosidase for		QY	1086	TTTGGGCGAGTTTTTCAAGCAGTATCTTGGCCCAATTAAGCTTAATATGATGCCAGTTGA	1145
CC	treating Gaucher's disease. Plants which are defective or deficient in		DB	1012	TTTGGGCGAGTTTTTCAAGCAGTATCTTGGCCCAATTAAGCTTAATATGATGCCAGTTGA	1071
CC	production of the GntI protein can be made to produce glycoproteins with		QY	1146	TTTGGGCGAGTTTTTCAAGCAGTATCTTGGCCCAATTAAGCTTAATATGATGCCAGTTGA	1205
CC	minimal, uniform and defined sugar residues, of low antigenicity. Use of		DB	1072	TTTGGGCGAGTTTTTCAAGCAGTATCTTGGCCCAATTAAGCTTAATATGATGCCAGTTGA	1131
CC	these plants eliminates the need for the difficult isolation and		QY	1206	CTTGGTTAAAAAGGCTTAAGCCCATCCATCGGAGCTGATGCTGCTTGAAGCAATTTAAAT	1265
CC	deglycosylation of native proteins or preparation in defective animal		DB	1132	CTTGGTTAAAAAGGCTTAAGCCCATCCATCGGAGCTGATGCTGCTTGAAGCAATTTAAAT	1191
CC	cells. This sequence represents potato (Solanum tuberosum) GntI encoding		QY	1266	AGATGGTATGTGCGTATTCAGTACAGATCACTAGACTTTGAAATATCGCAGGCA	1325
CC	cDNA		DB	1192	AGATGGTATGTGCGTATTCAGTACAGATCACTAGACTTTGAAATATCGCAGGCA	1251
XX						
SQ	Sequence 1669 BP; 489 A; 312 C; 387 G; 481 T; 0 U; 0 Other;		QY	1326	ATTTGGCATTTTTTGAAGAAATGGAAGATGGTGTACACGTGAGCATATAAAGGAATAGT	1385
	Query Match	72.3%; Score 1255; DB 2; Length 1669;	DB	1252	GTTCGGCATTTTTTGAAGAAATGGAAGATGGTGTACACGTGAGCATATAAAGGAATAGT	1311
	Best Local Similarity	89.1%; Pred. No. 0;				
	Matches 1412; Conservative	0; Mismatches 140; Indels 33; Gaps 4;	QY	1386	AGTTTTTCGGTACCAAAAGCTCCAGACGCTGATTTTCCTTGTGTGCCATGATTCGCTTCAACA	1445
QY	126	GATGAGAGGGAACAGTTTTTGTGTTGATTTCCGGTACCTCCATCTTGGCTGCTGCTGCG				185
DB	52	GATGAGAGGGAACAGTTTTTGTGTTGATTTTACGGTACCTTCGCTGCTGCTGCTGCG				111
QY	186	CTTCATCTACACACAGATCGGCTTTTTCGACACAGTGCAGATATGCAGATCGCCTTGC				245
DB	112	CTTCATCTACATACAGATCGGCTTTTTCGACACAGTGCAGATATGCAGATCGCCTTGC				171
QY	246	TGCTGCAATTTGAAGCAGAAAATCATTGTACAAGCCAGACAGATTCCTTTATTGACAGAT				305
DB	172	TGCTGCAATTTGAAGCAGAAAATCATTGTACAAGCCAGACAGATTCCTTTATTGACAGAT				231

*priority doc in instant appl.*

Db	1312	AGTTTTCCGGTTTCAACATCTAGACGTGTGTTCTTGTTCCTTCGATCTCTTCGACA	1371											PR	18-MAY-1999;	99US-0134768P.	
														PR	19-MAY-1999;	99US-0134941P.	
Qy	1446	ACTCGGAATTGAAGATACTTAAACAAAGATATGATTGCAGAGCCCGGGCAAAATTTTGA	1505											PR	20-MAY-1999;	99US-0135124P.	
														PR	21-MAY-1999;	99US-0135353P.	
Db	1372	ACTTGGAGTTGAAGATCTTAGCGAAGATATGATT---GGAGCCTGAGCAACAATTTAGA	1428											PR	24-MAY-1999;	99US-0135629P.	
														PR	25-MAY-1999;	99US-0136021P.	
Qy	1506	CTTATTGGTAGGATGATCGAGCTGACACTAAACCATGATTTT-----ACCAGTTAC	1558											PR	27-MAY-1999;	99US-0136392P.	
														PR	28-MAY-1999;	99US-0136782P.	
Db	1429	CTTATTGGTAGGATACATTTGAAGAGCTGACACGAAAGTATGACTACAGTAGTCTAC	1488											PR	01-JUN-1999;	99US-0137222P.	
														PR	01-JUN-1999;	99US-0137528P.	
Qy	1559	ATACAACGTTTTTAATGTATATACGGAGGAGCTCACTGTTCTAGTGTGAAGGATATCGGC	1618											PR	03-JUN-1999;	99US-0137502P.	
														PR	04-JUN-1999;	99US-0137502P.	
Db	1489	ATGCAACATTTTAATGTTAATGGAGGAAACCACTGCT-----T	1527											PR	07-JUN-1999;	99US-0137724P.	
														PR	08-JUN-1999;	99US-0138094P.	
Qy	1619	TTCTTAGTATTGGATGAATCATCAACAACTTATTTAAGTGTTCAGAACATAAAG	1678											PR	10-JUN-1999;	99US-0138847P.	
														PR	10-JUN-1999;	99US-0138847P.	
Db	1528	ATTGTTGGAATGGATGAATCATCACCACATCTCTATTATTTCAAGTTTACAAACATAAAG	1585											PR	14-JUN-1999;	99US-0139119P.	
														PR	16-JUN-1999;	99US-0139452P.	
Qy	1679	AGGAATGTAGCCCTGTTAAAGACTA	1703											PR	16-JUN-1999;	99US-0139453P.	
														PR	17-JUN-1999;	99US-0139453P.	
Db	1586	AGGAATGTTCGCCCTATAAAACAA	1610											PR	18-JUN-1999;	99US-0139457P.	
														PR	18-JUN-1999;	99US-0139457P.	
RESULT 3																	
AAC39540																	
ID	AAC39540 standard; DNA; 1639 BP.																
AC	AAC39540;																
XX	AC																
XX	17-OCT-2000 (first entry)																
XX	Arabidopsis thaliana																
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 25006.																
XX	XX																
KW	Hybridisation assay; genetic mapping; gene expression control;																
KW	protein identification; signal transduction pathway; metabolic pathway;																
KW	promoter; termination sequence; ss.																
XX	XX																
OS	Arabidopsis thaliana.																
XX	XX																
PN	EP1033405-A2.																
XX	XX																
PD	06-SEP-2000.																
XX	XX																
PF	25-FEB-2000; 2000EP-00301439.																
XX	XX																
PR	25-FEB-1999;																
PR	99US-0121825P.																
PR	05-MAR-1999;																
PR	99US-0123180P.																
PR	09-MAR-1999;																
PR	99US-0123548P.																
PR	23-MAR-1999;																
PR	99US-0125788P.																
PR	25-MAR-1999;																
PR	99US-0126264P.																
PR	29-MAR-1999;																
PR	99US-0126785P.																
PR	01-APR-1999;																
PR	99US-0127462P.																
PR	06-APR-1999;																
PR	99US-0128234P.																
PR	08-APR-1999;																
PR	99US-0128714P.																
PR	16-APR-1999;																
PR	99US-0129845P.																
PR	19-APR-1999;																
PR	99US-0130077P.																
PR	21-APR-1999;																
PR	99US-0130449P.																
PR	23-APR-1999;																
PR	99US-0130510P.																
PR	28-APR-1999;																
PR	99US-0130891P.																
PR	23-APR-1999;																
PR	99US-0131449P.																
PR	30-APR-1999;																
PR	99US-0132048P.																
PR	30-APR-1999;																
PR	99US-0132407P.																
PR	04-MAY-1999;																
PR	99US-0132484P.																
PR	04-MAY-1999;																
PR	99US-0132485P.																
PR	06-MAY-1999;																
PR	99US-0132486P.																
PR	06-MAY-1999;																
PR	99US-0132487P.																
PR	07-MAY-1999;																
PR	99US-0132863P.																
PR	11-MAY-1999;																
PR	99US-0132456P.																
PR	14-MAY-1999;																
PR	99US-0134218P.																
PR	14-MAY-1999;																
PR	99US-0134219P.																
PR	14-MAY-1999;																
PR	99US-0134221P.																
PR	14-MAY-1999;																
PR	99US-0134370P.																
PR	14-MAY-1999;																
PR	99US-0145913P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																
PR	99US-0145918P.																

PR	27-JUL-1999;	99US-0145919P.	PR	26-OCT-1999;	99US-0161359P.
PR	28-JUL-1999;	99US-0145919P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146388P.	PR	28-OCT-1999;	99US-0161920P.
PR	02-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-0161920P.
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161992P.
PR	04-AUG-1999;	99US-0147204P.	PR	28-OCT-1999;	99US-0161993P.
PR	04-AUG-1999;	99US-0147302P.	PR	29-OCT-1999;	99US-0162142P.
PR	05-AUG-1999;	99US-0147192P.	Query Match 44.4%; Score 770.4; DB 3; Length 1639;		
PR	05-AUG-1999;	99US-0147260P.	Best Local Similarity 73.2%; Pred. No. 9e-224;		
PR	06-AUG-1999;	99US-0147303P.	Matches 987; Conservative 0; Mismatches 361; Indels 0; Gaps 0;		
PR	06-AUG-1999;	99US-0147416P.			
PR	09-AUG-1999;	99US-0147493P.			
PR	09-AUG-1999;	99US-0147935P.			
PR	10-AUG-1999;	99US-0148171P.			
PR	11-AUG-1999;	99US-0148319P.			
PR	12-AUG-1999;	99US-0148341P.			
PR	13-AUG-1999;	99US-0148565P.			
PR	13-AUG-1999;	99US-0148684P.			
PR	16-AUG-1999;	99US-0149368P.			
PR	17-AUG-1999;	99US-0149175P.			
PR	18-AUG-1999;	99US-0149426P.			
PR	20-AUG-1999;	99US-0149722P.			
PR	20-AUG-1999;	99US-0149723P.			
PR	20-AUG-1999;	99US-0149929P.			
PR	23-AUG-1999;	99US-0149902P.			
PR	23-AUG-1999;	99US-0149930P.			
PR	25-AUG-1999;	99US-0150566P.			
PR	26-AUG-1999;	99US-0150884P.			
PR	27-AUG-1999;	99US-0151065P.			
PR	27-AUG-1999;	99US-0151066P.			
PR	27-AUG-1999;	99US-0151080P.			
PR	30-AUG-1999;	99US-0151303P.			
PR	31-AUG-1999;	99US-0151438P.			
PR	01-SEP-1999;	99US-0151930P.			
PR	07-SEP-1999;	99US-0152363P.			
PR	10-SEP-1999;	99US-0153070P.			
PR	13-SEP-1999;	99US-0153758P.			
PR	15-SEP-1999;	99US-0154018P.			
PR	16-SEP-1999;	99US-0154039P.			
PR	20-SEP-1999;	99US-0154779P.			
PR	22-SEP-1999;	99US-0155139P.			
PR	23-SEP-1999;	99US-0155486P.			
PR	24-SEP-1999;	99US-0155659P.			
PR	28-SEP-1999;	99US-0156458P.			
PR	29-SEP-1999;	99US-0156596P.			
PR	04-OCT-1999;	99US-0157117P.			
PR	05-OCT-1999;	99US-0157539P.			
PR	06-OCT-1999;	99US-0157865P.			
PR	07-OCT-1999;	99US-0158029P.			
PR	08-OCT-1999;	99US-0158232P.			
PR	12-OCT-1999;	99US-0158369P.			
PR	13-OCT-1999;	99US-0159293P.			
PR	13-OCT-1999;	99US-0159294P.			
PR	13-OCT-1999;	99US-0159295P.			
PR	14-OCT-1999;	99US-0159329P.			
PR	14-OCT-1999;	99US-0159330P.			
PR	14-OCT-1999;	99US-0159331P.			
PR	14-OCT-1999;	99US-0159637P.			
PR	14-OCT-1999;	99US-0159638P.			
PR	18-OCT-1999;	99US-0159584P.			
PR	21-OCT-1999;	99US-0160741P.			
PR	21-OCT-1999;	99US-0160767P.			
PR	21-OCT-1999;	99US-0160768P.			
PR	21-OCT-1999;	99US-0160770P.			
PR	21-OCT-1999;	99US-0160814P.			
PR	21-OCT-1999;	99US-0160815P.			
PR	22-OCT-1999;	99US-0160980P.			
PR	22-OCT-1999;	99US-0160981P.			
PR	22-OCT-1999;	99US-0160989P.			
PR	25-OCT-1999;	99US-0161404P.			
PR	25-OCT-1999;	99US-0161405P.			
PR	25-OCT-1999;	99US-0161406P.			
PR	27-JUL-1999;	99US-0145919P.	PR	26-OCT-1999;	99US-0161359P.
PR	28-JUL-1999;	99US-0145919P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146388P.	PR	28-OCT-1999;	99US-0161920P.
PR	02-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-0161920P.
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161992P.
PR	04-AUG-1999;	99US-0147204P.	PR	28-OCT-1999;	99US-0161993P.
PR	04-AUG-1999;	99US-0147302P.	PR	29-OCT-1999;	99US-0162142P.
PR	05-AUG-1999;	99US-0147192P.	Query Match 44.4%; Score 770.4; DB 3; Length 1639;		
PR	05-AUG-1999;	99US-0147260P.	Best Local Similarity 73.2%; Pred. No. 9e-224;		
PR	06-AUG-1999;	99US-0147303P.	Matches 987; Conservative 0; Mismatches 361; Indels 0; Gaps 0;		
PR	06-AUG-1999;	99US-0147416P.			
PR	09-AUG-1999;	99US-0147493P.			
PR	09-AUG-1999;	99US-0147935P.			
PR	10-AUG-1999;	99US-0148171P.			
PR	11-AUG-1999;	99US-0148319P.			
PR	12-AUG-1999;	99US-0148341P.			
PR	13-AUG-1999;	99US-0148565P.			
PR	13-AUG-1999;	99US-0148684P.			
PR	16-AUG-1999;	99US-0149368P.			
PR	17-AUG-1999;	99US-0149175P.			
PR	18-AUG-1999;	99US-0149426P.			
PR	20-AUG-1999;	99US-0149722P.			
PR	20-AUG-1999;	99US-0149723P.			
PR	20-AUG-1999;	99US-0149929P.			
PR	23-AUG-1999;	99US-0149902P.			
PR	23-AUG-1999;	99US-0149930P.			
PR	25-AUG-1999;	99US-0150566P.			
PR	26-AUG-1999;	99US-0150884P.			
PR	27-AUG-1999;	99US-0151065P.			
PR	27-AUG-1999;	99US-0151066P.			
PR	27-AUG-1999;	99US-0151080P.			
PR	30-AUG-1999;	99US-0151303P.			
PR	31-AUG-1999;	99US-0151438P.			
PR	01-SEP-1999;	99US-0151930P.			
PR	07-SEP-1999;	99US-0152363P.			
PR	10-SEP-1999;	99US-0153070P.			
PR	13-SEP-1999;	99US-0153758P.			
PR	15-SEP-1999;	99US-0154018P.			
PR	16-SEP-1999;	99US-0154039P.			
PR	20-SEP-1999;	99US-0154779P.			
PR	22-SEP-1999;	99US-0155139P.			
PR	23-SEP-1999;	99US-0155486P.			
PR	24-SEP-1999;	99US-0155659P.			
PR	28-SEP-1999;	99US-0156458P.			
PR	29-SEP-1999;	99US-0156596P.			
PR	04-OCT-1999;	99US-0157117P.			
PR	05-OCT-1999;	99US-0157539P.			
PR	06-OCT-1999;	99US-0157865P.			
PR	07-OCT-1999;	99US-0158029P.			
PR	08-OCT-1999;	99US-0158232P.			
PR	12-OCT-1999;	99US-0158369P.			
PR	13-OCT-1999;	99US-0159293P.			
PR	13-OCT-1999;	99US-0159294P.			
PR	13-OCT-1999;	99US-0159295P.			
PR	14-OCT-1999;	99US-0159329P.			
PR	14-OCT-1999;	99US-0159330P.			
PR	14-OCT-1999;	99US-0159331P.			
PR	14-OCT-1999;	99US-0159637P.			
PR	14-OCT-1999;	99US-0159638P.			
PR	18-OCT-1999;	99US-0159584P.			
PR	21-OCT-1999;	99US-0160741P.			
PR	21-OCT-1999;	99US-0160767P.			
PR	21-OCT-1999;	99US-0160768P.			
PR	21-OCT-1999;	99US-0160770P.			
PR	21-OCT-1999;	99US-0160814P.			
PR	21-OCT-1999;	99US-0160815P.			
PR	22-OCT-1999;	99US-0160980P.			
PR	22-OCT-1999;	99US-0160981P.			
PR	22-OCT-1999;	99US-0160989P.			
PR	25-OCT-1999;	99US-0161404P.			
PR	25-OCT-1999;	99US-0161405P.			
PR	25-OCT-1999;	99US-0161406P.			

Db	938	CGACGGAGTCTGTAGAACATACAAATTTGGTGAACATGGTCTAGTTTGGGACATTT	997	PR	06-MAY-1999;	99US-0132486P.
Qy	1099	TTCAAGCAGTATCTTGAGCCAAATTAACATAATGATGTCAGGTTGATTCGAAGTCAATG	1158	PR	06-MAY-1999;	99US-0132487P.
Db	998	TTCAAGCAGTATCTTGAGCCAAATTAACATAATGATGTCAGGTTGATTCGAAGTCAATG	1057	PR	07-MAY-1999;	99US-0132863P.
Qy	1159	GACCTTAGTTACCTTTTGGAGGACAAATTAACGTGAAACACTTTGGTGACCTTGGTTAAAG	1218	PR	11-MAY-1999;	99US-0134256P.
Db	1058	GACCTGGATACCTTGACAGAGGGAAACTATACCAAGTACTTTCTGGCTTAGTGAGACAA	1117	PR	14-MAY-1999;	99US-0134218P.
Qy	1219	GCTAAGCCCATCCATGGAGCTGATGCTGTCTTTGAAAGCATTTAACATAGATGGTGAATG	1278	PR	14-MAY-1999;	99US-0134219P.
Db	1118	GCACGACCAATTCAGAGTTCTTGACCTTGCTTTAAAGGCTCAAAACATAAAGGATGATTT	1177	PR	14-MAY-1999;	99US-0134221P.
Qy	1279	CGTATTCAGTACAGAGATCAACTAGACTTTGAAATATCGCACGGCAATTTGGCATTTTT	1338	PR	18-MAY-1999;	99US-0134370P.
Db	1178	CGTATCCGGTATAAAGACCAAGTAGAGTTTGAACCGCATTCGACGGGAATTTGGTATATTT	1237	PR	18-MAY-1999;	99US-0134768P.
Qy	1339	GAAGAAATGGAAGGATGGTGTACCAAGTGCAGCATATAAAGGAATAGTTTCCGGTAC	1398	PR	19-MAY-1999;	99US-0134941P.
Db	1238	GAAGAAATGGAAGGATGGTGTCCCTCGAACACATATAAAGGATAGTGGTTCGAAATC	1297	PR	20-MAY-1999;	99US-0135124P.
Qy	1399	CAACAGCTCCAGACGCTGATTTCTTTGGCCCATGATTCGCTTCAACAACTCGGAATTGAA	1458	PR	21-MAY-1999;	99US-0135353P.
Db	1298	CAGACAACACGCTGATTTCTTGTTGGGCCAGATTCGTAAATGCAGCTTGGAAATCGA	1357	PR	24-MAY-1999;	99US-0135629P.
Qy	1459	GATACTTAAACAAAGATGATGTCAGGA	1486	PR	25-MAY-1999;	99US-0136021P.
Db	1358	AATTCCTGATCAAAACATATGAAAGGA	1385	PR	28-MAY-1999;	99US-0136392P.
RESULT 4						99US-0136782P.
ID	AAC50197 standard; DNA; 1641 BP.					99US-0137222P.
XX	AAC50197;					99US-0137528P.
AC	AAC50197;					99US-0137502P.
XX	18-OCT-2000 (first entry)					99US-0137724P.
DT	Arabidopsis thaliana DNA fragment SEQ ID NO: 63936.					99US-0138094P.
DE	Hybridisation assay; genetic mapping; gene expression control;					99US-0138540P.
XX	protein identification; signal transduction pathway; metabolic pathway;					99US-0139119P.
KW	promoter; termination sequence; ss.					99US-0139452P.
XX	Arabidopsis thaliana.					99US-0139453P.
OS	EP1033405-A2.					99US-0139454P.
XX	06-SEP-2000.					99US-0139455P.
XX	25-FEB-2000; 2000EP-00301439.					99US-0139456P.
PF	25-FEB-1999;					99US-0139457P.
XX	05-MAR-1999;					99US-0139458P.
PR	09-MAR-1999;					99US-0139460P.
PR	23-MAR-1999;					99US-0139461P.
PR	25-MAR-1999;					99US-0139462P.
PR	01-APR-1999;					99US-0139463P.
PR	06-APR-1999;					99US-0139750P.
PR	16-APR-1999;					99US-0139763P.
PR	19-APR-1999;					99US-0139817P.
PR	23-APR-1999;					99US-0139899P.
PR	28-APR-1999;					99US-0140354P.
PR	30-APR-1999;					99US-0140695P.
PR	04-MAY-1999;					99US-0140823P.
PR	05-MAY-1999;					99US-0140991P.



Db 816 TTTCTGCGCTTGGTGGATGCTCAAGAGATCGACTGGGATGATGTTATACCAAAATGG 875  
Qy 976 CCAAAGG-CTTACTGGGACGACTGGCTAAGACTCAAGAGAAATCACAGAGTGCACAATT 1034  
Db 876 CCAAAGGCTTTACTGGGATGATTGGCTGAGACTAAAGGAAACCAATAAAGCGCCCAATT 935  
Qy 1035 TATTGCGCCAGAAAGTTTCAGAACATATAATTTGGTGGAGCATGTTCTAGTTTGGGGCA 1094  
Db 936 CATTCGACCGGAAGTCTGTAGAACATACAATTTTGGTGAACATGGGTCTAGTTTGGGACA 995  
Qy 1095 GTTTTTCAGGAGTATCTTGAGCCAAATTAATACTAAATGATGCTCCAGGTGATTGGAAGTC 1154  
Db 996 GTTTTTCAGTCACTGAGAACCTTAAGACTAAACGATGAGCGTTGACTGGAAAGC 1055  
Qy 1155 AATGACCTTAGTTACCTTTTGGAGGACAAATACGTGAAACACTTTGGTGACTTGGTTAA 1214  
Db 1056 AAGGACCTGGGATACCTGACAGAGGGAAACTATACCAAGTACTTTTCTGGCTTAGTGAG 1115  
Qy 1215 AAGGCTAAGCCCATCCATGAGCTGATGCTGCTTGAAGCATTTAACAATAGATGTGA 1274  
Db 1116 ACNAGCAGCAATCAAGGTTCTGACCTTGTCTTAAAGGCTCAAAACATATAAGSATGA 1175  
Qy 1275 TGTGCGTATTCACTACAGAGATCAACTAGACTTTTGAATAATCGACGGCAATTTGGCAT 1334  
Db 1176 TGTTCGATCCGTTATAAAGCAAGTAGATGTTTGAACGCAATTCAGGGNAITTTGGTAT 1235  
Qy 1335 TTTTGAAGAATGGAAGGATGTTGACCAAGTGCAGCATATAAAGAAATAGTATTTCCG 1394  
Db 1236 ATTTGAAGAATGGAAGGATGTTGCTTCGAAACAGCATATAAAGGATAGTGTGTTTCG 1295  
Qy 1395 GTACCAACGTCAGACGCTGATTCCTTGTGTCGCAATGATTCGCTTCAACACTCGGAAT 1454  
Db 1296 AATCCAGACAAACAGACGCTGATTCCTGTTGGTGGCCAGATTCTGTAATGCAGCTTGAAT 1355  
Qy 1455 TGAAGATACTTAACAAAGATATATGTCAGGA 1486  
Db 1356 TCGAAATTCCTGATGCAAAACATATGAAGGA 1387

RESULT 5  
ADXS4592  
ID ADXS4592 standard; cDNA; 1836 BP.  
XX  
AC ADXS4592;  
XX  
DT 21-APR-2005 (first entry)  
XX  
DE Plant full length insert polynucleotide seqid 29332.  
XX  
KW plant protectant; plant growth regulant; gene therapy; plant;  
KW recombinant DNA construct; physical array; plant breeding marker;  
KW cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;  
KW extreme osmotic condition; pathogen tolerance; pest tolerance;  
KW growth rate; cell cycle pathway; disease resistance;  
KW galactomannan production; lignin production; plant growth regulator;  
KW yield; plant growth; plant development; seed oil; protein yield;  
XX protein content; gene; ss.  
OS Unidentified.  
XX  
XX US2004034888-A1.  
XX  
PD 19-FEB-2004.  
XX  
XX 28-APR-2003; 2003US-00425114.  
XX  
XX 06-MAY-1999; 99US-00304517.  
XX  
XX 05-NOV-2001; 2001US-00985678.  
XX  
XX (LIUJ/) LIU J.  
XX (ZHOU/) ZHOU Y.  
XX (KOVA/) KOVALIC D K.

PA (SCRE/) SCREEN S E.  
PA (TABA/) TABASKA J E.  
PA (CAOY/) CAO Y.  
XX  
XX Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;  
XX WPI; 2004-180133/17.  
XX  
XX New recombinant DNA construct, useful for improving plant tolerance to  
XX cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or  
XX pests, for conferring increased resistance to plant disease, or for  
XX improving yield.  
XX  
XX Claim 1; SEQ ID NO 29332; 15pp; English.  
XX  
XX The invention describes a recombinant DNA construct comprising a  
XX polynucleotide consisting of a sequence encoding an amino acid sequence  
XX available in electronic form from the US patent office at  
XX ftp.seqdata.uspto.gov/sequence.html?DocID:2004034888. The polynucleotide  
XX of the invention are also useful in physical arrays of molecules and as  
XX plant breeding markers. The recombinant DNA construct is useful for  
XX improving plant tolerance to cold, heat, drought, herbicides, extreme  
XX osmotic conditions, pathogens or pests, for manipulating growth rate in  
XX plant cells by modification of the cell cycle pathway, for conferring  
XX increased resistance to plant disease, for producing galactomannan,  
XX lignin or plant growth regulators, for increasing the rate of homologous  
XX recombination in plants, for improving yield by modification of  
XX photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake  
XX or by providing improved plant growth and development under at least one  
XX stress condition or for modifying seed oil or protein yield and/or  
XX content. This sequence represents a plant full length insert  
XX polynucleotide that can be used in the recombinant DNA construct of the  
XX invention.  
XX  
SQ Sequence 1836 BP; 546 A; 326 C; 406 G; 558 T; 0 U; 0 Other;  
Query Match 41.1%; Score 714.2; DB 13; Length 1836;  
Best Local Similarity 71.8%; Pred. NO. 1.4e-206;  
Matches 1037; Conservative 0; Mismatches 288; Indels 120; Gaps 3;  
Qy 139 AAGTTTGTGCTGATTTTCGGTACCTCCTCATCTTGGCTGCTGCTTCATCTACACA 198  
Db 43 AAAGTTTCTGTGATTTTCGGTTTCTCTCTGATTCGAGCTGGCTTCATCTATATC 102  
Qy 199 CAGATGCGGCTTTTTCGGACACAGTGCAGATATGCAGATGCGCTTGCCTGCAATGAA 258  
Db 103 CAGATGAGGCTTTTTCGCAACACAAATCAGAATATGCTGCTGCTGCTGATTCGAA 162  
Qy 259 GCAGAAATCATGTACAGCCAGACAGATTCCTTATTGACAGATTAGCCTGCAGCAA 318  
Db 163 GCTGAAACCAATGTACAAAGTCAACCGCATCGCTGATTCAGATTAGCTTGCAACA 222  
Qy 319 GGAAGAAATAGTTGCTCTCTTGAAGAACAAAT----- 347  
Db 223 GGAGAAATTTGCGCCCTGGAAGATTAAATTTGGTCTGTTCTTTCAGAGAGC 282  
Qy 348 -GAAGCGTCAGGACACAGGAGTGCAGCAATTAAGGCTCTTGTTCAGGATCTTGAAGTA 406  
Db 283 AGAAACGTCGAGACCAAGAATTTGGACAAACAAAGTCTCTTTGACAAAGATCTTTGAAAGAA 342  
Qy 407 AGGCGATAAAGATTTGATCGGAAATGTACAGATGCGAGTGGCTGCTGTAGTTGTATGG 466  
Db 343 AAGACCTGCAGAGCTGATTTGATAAGTGCAGGTTCCGGTGGCTGCTTGTGTGATCAGG 402  
Qy 467 CTTGCAATCGGCTGATTTACCTGGAAAAAGACTATAATCCATCTTTAAAAATACCAATAT 526  
Db 403 CATGTAATCGTCTGATTACCTGGAGGACTATTAATCTGATTGAAATACCAAGGC 462  
Qy 527 CTGTTGCGTCAAAATATCTCTTTTCATATCCAGGATGATCATCTCTGATGTGACGA 586  
Db 463 CCATTTCTTCAGATATCTCTTTTATTTGTTATCTCAGGATGATCAAAATCCAAATGTTAAA 522  
Qy 587 AGCTTGTCTTTGAGCTATGATCAGCTGACGTTATATGACAGACTTTGGATTTTGAACCTGTGC 646

523	GTAAAGGCTTTGAGCTATGATCAGTTATCTTATATGAGCAGCTTGGATTTTGAACCACTTC	582	XX	21-APR-2005 (first entry)
647	ATACTGAAAGACCAAGGAGCTGATGTCATACATACTACAAAATTCGACGTCATTTACAAGTGGG	706	XX	Plant full length insert polynucleotide seqid 487.
583	AAACTGAAAGCCCTGGAGAGTTAACTGCTTACTACAAAATTCGACGTCATTTACAAGTGGG	642	XX	plant protectant; plant growth regulant; gene therapy; plant;
707	CATTGGATCAGCTGTTTTCACAAGCATAAATTTAGCCGTGTTTATCATACTAGAAATGATA	766	XX	recombinant DNA construct; physical array; plant breeding marker;
643	CTCTAGATCAACTGTTCTTACAAGCATAAATCTTCAGCCGTGTGATCATCTTTGAAGATGACA	702	XX	cold tolerance; heat tolerance; drought tolerance; herbicide tolerance;
767	TGAAAATTTGCCCTTGATTTTTCGACTTTTGTGAGGCTGGAGCTACTCTTCTTGACAGAG	826	XX	extreme osmotic condition; pathogen tolerance; pest tolerance;
703	TGAAAATAGCACCTGATTTCTTGATTTTGAAGCTGCAGCGACTCTCTCTGACAAGG	762	XX	growth rate; cell cycle pathway; disease resistance;
827	ACAAAGTGAATATGGCTATTTCTTCTTGGAATGACAATGGACAAATGACAGTTTGTCCAAG	886	XX	galactomannan production; lignin production; plant growth regulator;
763	ATAAATCCATTTATGGCTGTTTCTCATGGAATGACAATGGACAAAAGCAGTTTGTACATG	822	XX	yield; plant growth; plant development; seed oil; protein yield;
887	ATCCTTATGCTCTTTACCGCTCAGATTTTTCCTCCGGTCTTGGATGGATGCTTTCAAAAT	946	OS	protein content; gene; ss.
823	ATCCATATGAACCTTTATGCTCAGACTTCTTTCTCTGGATTAGGATGGATGCTGGCCAGAT	882	XX	Unidentified.
947	CTACTTGGGACGAATTTATCTCAAAGTGGCCAAAGGCTTACTGGGACGACTGGCTTAAGAC	1006	XX	US2004034888-A1.
883	CTACATGGATGAGCTATCACCAAAATGGCCAAAAG-----CTGGGATGACTGGTTGAGAC	938	XX	19-FEB-2004.
1007	TCAAAGAGAATCAGAGGTCGACAAATTTATTCGCCCGAGAAAGTTTGCAGAACATATAAT	1066	XX	28-APR-2003; 2003US-00425114.
939	TTAAAGAGAATCAAAAGGACGACAGTTTATCCGGCCGAAAGTATGCAAGACATATAAT	998	XX	06-MAY-1999; 99US-00304517.
1067	TTGGTGACATGTTCTAGTTTGGGCGAGTTTTCACAGCATATCTTCAGGCCAATTAAC	1126	XX	05-NOV-2001; 2001US-00985678.
999	TTGGTGACATGTTCTAGTTTGGGCGAGTTTTCACAGCATATCTTCAGGCCAATTAAC	1058	XX	(LIUJ/) LIU J.
1127	TAAATGATGTC-----	1137	XX	(ZHOU/) ZHOU Y.
1059	TGAATGATGTCAAGGTACACCTGCTTTTATGAACAATACAGCTCTTAATATGTTCTTAT	1118	XX	(KOVA/) KOVALIC D K.
1138	-----CAGTTGATTTGGAAGTCAATGGACC	1162	XX	(SCRE/) SCREEN S E.
1119	TATTAGAATGATGCTATTTCACTTTGGATTTTGGAGGTTGATTTGGAATCAATGGATT	1178	XX	(TABAJ/) TABASKA J E.
1163	TTAGTTACTTTTGGAGACAAATACGTTTGGTGAACACTTTGGTGAACCTTAAAGGCTA	1222	XX	(CAOY/) CAO Y.
1179	TGAGCTATTTTACTGGAGGATAAATATTTATGCACTTTTGGCAACGTTGTTAAGAAAGCTA	1238	XX	Liu J, Zhou Y, Kovalic DK, Screen SE, Tabaska JE, Cao Y;
1223	AGCCCATCCATGGAGCTGATGCTGTTTGAAGCAATTTAAACATAGATGCTGATGCGTA	1282	XX	WPI; 2004-180133/17.
1239	CACCTGCTCTATGGAGCTGACATGGTTCTTAAAGGCATATAATATAGATGGCATGTCGTA	1298	XX	New recombinant DNA construct, useful for improving plant tolerance to
1283	TTCACTCAGAGATCAACTAGACTTTGAAAATATCGCAGCGCAATTTGGCATTTTGAAG	1342	XX	cold, heat, drought, herbicides, extreme osmotic conditions, pathogens or
1299	TCAAAATGAAGATCAGTCAGACTTTGAAAACATTTGCTCGCAAATTTGGTATATTTCAAG	1358	XX	pests, for conferring increased resistance to plant disease, or for
1343	AATGAAGGATGGTGTACCAAGTCGAGCATATAAAGGAATAGTATGTTTCCGGTACCAA	1402	XX	improving yield.
1359	AGTGAAGGATGGTGTGCCAGGACAGCCTATAAAGGAGTAGTCTGTTTTCAGATATCAA	1418	XX	Claim 1; SEQ ID NO 487; 15pp; English.
1403	CGTCAGAGCTGATTTCTTGTGTCATGATTTGCTTCAACAACCTCGGAATTTGAAGATA	1462	XX	The invention describes a recombinant DNA construct comprising a
1419	CCTCAGAGCTATATTTCTTGTGTCAGATATTTTGAAGCTACTTTCAGATCGAAGAT	1478	XX	polynucleotide consisting of a sequence encoding an amino acid sequence
1463	CTTAA 1467		XX	available in electronic form from the US patent office at
1479	CTTAA 1483		XX	ftp.secdatala.uspto.gov/sequence.html?docID:2004034888. The polynucleotide
RESULT 6			XX	of the invention are also useful in physical arrays of molecules and as
ADO81767			XX	plant breeding markers. The recombinant DNA construct is useful for
ID			XX	improving plant tolerance to cold, heat, drought, herbicides, extreme
XX			XX	osmotic conditions, pathogens or pests, for manipulating growth rate in
AC			XX	plant cells by modification of the cell cycle pathway, for conferring
			XX	increased resistance to plant disease, for producing galactomannan,
			XX	lignin or plant growth regulators, for increasing the rate of homologous
			XX	recombination in plants, for improving yield by modification of
			XX	photosynthesis or carbohydrate, nitrogen or phosphorus use and/or uptake
			XX	or by providing improved plant growth and development under at least one
			XX	stress condition or for modifying seed oil or protein yield and/or
			XX	content. This sequence represents a plant full length insert
			XX	polynucleotide that can be used in the recombinant DNA construct of the
			XX	invention.
			XX	Sequence 1882 BP; 530 A; 363 C; 473 G; 516 T; 0 U; 0 Other;
			XX	Query Match 38.1%; Score 661.8; DB 13; Length 1882;
			XX	Best Local Similarity 68.4%; Pred. NO. 1.4e-190;
			XX	Matches 931; Conservative 0; Mismatches 427; Indels 3; Gaps 1;
			QY	147 CTGTGATTTCCGGTACCTCCTCATCTTGGCTGCTGCTGCTTCATCTACACAGATGCG 206
			DB	160 CTGCGACCTTCGCGCTTCCTCCTCCTCGTCGACGTGACGCTTCATCTACATCAAGTGGG 219
			QY	207 GCTTTTTGCGACACAGTCAGATATGCGATCGCCTTGTCTGCTCAATTTGAAGCAGAAAA 266

Db 220 CCTCTTGTCTACTCAGTCCCATATATGCGGCGCTCTTGGGAGCAGAAAGATCTGAAA 279  
Qy 267 TCATGTACAGCAGCAGACAGATTGCTTTATGACCAAGATTAGCCTGCGAGCAAGAAAT 326  
Db 280 TCAATGCACCAAGTCAATTTAAAGTCCCTTGATCGATCAAGTCAGCATGCGAGGAGAAGAT 339  
Qy 327 AGTTGCTCTTGAGAACCAATGAAGCTCAGGCCAGGAGTGGCGACAATTAAAGGCTCT 386  
Db 340 TGTAGCAATTGGAAGAGATAAAGGTACGGCAAGATGAAGAACGTGCACATCTGAGGATTTT 399  
Qy 387 TGTTCAGGATCTTGAAAGTAAAGGCATAAAAAGTTGATCGGAAATGTACAGATGCCAGT 446  
Db 400 GATAAAGGATCTTGAAAAAGGAGTGTGCAAGACTACTAGACAGAAATGTGGTCCCGT 459  
Qy 447 GGCTGCTGTAGTTTATGCTTGCAATCGGGCTGATTAACCTGGAAAAAGACTATTAAATC 506  
Db 460 TGCTGCTGTGTCTATAATGCTTGCAATCGACCAAGACTATTTAGAGAGAAACAGTTGAATC 519  
Qy 507 CATCTTAAATACCAATATCTGTGTGCTCAAAATATCCTCTTTTATATCTCAGGATGG 566  
Db 520 TATCCTGAAGTATCAGACATCAGTTGCTTCGAAAGTTCCGCTTTTATATCTCAGGATGG 579  
Qy 567 ATCATCTCTGATGTCAGGAAGCTTGCTTTGAGCTATGATCAGCTGACGTATATGCAGCA 626  
Db 580 AGCAATGGAGCTGTGAAGATAAAGCTTTGGATATAAACAATAACATACATGACGCA 639  
Qy 627 CTTGGATTTTGAACCTGTGCATCTGAAAGACCAAGGGAGCTGATTCATACATAACAAAT 686  
Db 640 TGTGGATCTTGAACCTGTCAAACTGAAAGCCAGGAGAAATTGACAGCATATTACAAGAT 699  
Qy 687 TGCAGCTCANTACNAAGTGGGATGGATCAGCTGTTTACAAGCATAAATTTAGCCGTGT 746  
Db 700 TGCTAAACATTTATAAGTGGGCTTGGACAACCTATTCAATTAACATAAATTTGCTCGAT 759  
Qy 747 TATCATACTAGAACATGATGGAATTTGCCCTGATTTTGTGACTTTTTCAGGCTGG 806  
Db 760 AATCATTTCTAGAGATGACATGGAGATTGCCCGGATTTTTCGACTACTTCGAGGCTGC 819  
Qy 807 AGCTACTCTTTTGACAGACAGATCGATTATGGCTTAATTTCTTTGGAATGACAATGG 866  
Db 820 AGCTAAACTACTTGATTAATGACAAGACGATATGGCTGTTTTCGCTTTGGAATGACAATGG 879  
Qy 867 ACAATGCGATTGTCCAAGATCCTTATGCTCTTTACCGCTCAGATTTTTCGCGTCT 926  
Db 880 GCAGAAGCAGTTGTAAACGACCCAAAGCTCTTTACCGCTCAGATTTCTTCTCGGCT 939  
Qy 927 TGGATGATGCTTTCAAAATCTACTTGGGAGCAATTAATCTCCAAAGTGGCCAAAGGCTTA 986  
Db 940 TGGATGATGTTAAACAAGTCCACTTGGATTTGAATTTGTACCAAAAGTGGCTTAAGCTTA 999  
Qy 987 CTGGACGCTGGCTTAAGACTCAAGAGAAATCAAGAGGTCGACAATTTATTCGCCAGA 1046  
Db 1000 TTGGGATGATTTGGGTGAGGCTTAAGGAGGTACATGGACATCGGCAATTCATACGCCGGA 1059  
Qy 1047 AGTTTGACAGACATATAATTTTGTGTAGCAGTGTCTAGTTTGGGCGAGTTTTCAGCA 1106  
Db 1060 AATTTGAGAACATACAAATTTTGGCAAGCATGGATCAAGCTTGGACAGTTCTTCGAGCA 1119  
Qy 1107 GTATCTTTGAGCCAATTAACATTAATGATGTCCAGTTTGAATTTGGAAGTCAATGGACCTTAG 1166  
Db 1120 ATACTTGGAGCCCATTAAGTTTAAACGATTTCAATCGATTTGGAATTTCTGAGGATCTGAG 1179  
Qy 1167 TTACCTTTTGGAGGACAATTAACGTAACACATTTTGGTGACTTGGTTAAAAAGGCTAAGCC 1226  
Db 1180 CTACCTTTAGGAGGATTAAGTTTCAACAAATTTGGAAGAGAGGTGGCTAGTGCCACTCC 1239  
Qy 1227 CATCCATGGAGCTGATCTGTCTTTGAAAGCAATTTAAACATAGATGGTGTGCGTATTCA 1286  
Db 1240 TCTCCGTTGAGTACGATCTGTGTTGAAAGGCCCAATATGGCTGAGGATGTAAAGGATCCA 1299  
Qy 1287 GTACAGAGATCAACTAGACTTTTGAATAATATCGCACGGCAATTTGGCACTTTTGAAGATG 1346

Db 1300 ATACGACGACCAAGAGGTTTTCGAGCAGATAGCTGCTCAATTTGGAATATTGCAAGAATG 1359  
Qy 1347 GAAGGATGGTGTACCAAGTGCAGCATATAAAGAAATAGTAGTTTTCGGGTACCA---AAC 1403  
Db 1360 GAAGGATGGCATCCCAAGAACGGCTTTACAAAGGAGTGGTGGTCTTCGGGTACCAAGTAG 1419  
Qy 1404 GTCCAGACGTGATTTCTTGTGGCCATGATTCGCTTCAACCACTCGGAATTCGAAGATAC 1463  
Db 1420 TCAAGGCGAATATTTCTCGTAGCCAGATTTCTTCGTGCTGGGGGTGAGCATGT 1479  
Qy 1464 TTAACAAAATATGATGATTCAGGAGCCCGGCGCAAAATTTTGG 1504  
Db 1480 TGAGAGGTAGATTTTACCCCTTGGATGGCAATATCATG 1520

RESULT 7  
ADR61718  
ID ADR61718 standard; cDNA; 1211 BP.  
XX  
AC ADR61718;  
XX  
DT 02-DEC-2004 (first entry)  
XX  
Cotton cDNA sequence, SEQ ID 2499.  
DE  
XX  
KW Cotton; ss; plant; cold tolerance; growth rate; cell cycle pathway;  
KW drought tolerance; plant disease resistance; galactomannan; lignin;  
KW plant growth regulator; heat tolerance; herbicide tolerance;  
KW homologous recombination; extreme osmotic condition tolerance;  
KW pathogen resistance; pest resistance; yield; photosynthesis; seed oil;  
KW stress resistance.  
XX  
OS Gossypium hirsutum.  
XX  
PN US2004181830-A1.  
XX  
PD 16-SEP-2004.  
XX  
PF 29-JAN-2004; 2004US-00767795.  
XX  
PR 07-MAY-2001; 2001US-00849529.  
PR 12-DEC-2001; 2001US-00021323.  
XX  
XX (KOVA/) KOVALIC D K.  
PA (ZHOU/) ZHOU Y.  
PA (CAOY/) CAO Y.  
XX  
PI Kovalic DK, Zhou Y, Cao Y;  
XX  
DR WPI; 2004-667718/65.  
XX  
PT New recombinant nucleic acid molecules and polypeptides from Gossypium  
PT hirsutum, useful for producing plants with improved biological  
PT characteristics (e.g. improved plant cold or drought tolerance).  
XX  
PS Claim 1; SEQ ID NO 2499; 14pp; English.  
XX  
CC The invention relates to a recombinant polynucleotide comprising any of  
CC the 58798 Cotton plant cDNA sequences mentioned in the specification.  
CC Also a recombinant polypeptide comprising any of the 58798 amino acid  
CC sequences mentioned in the specification and producing a plant having an  
CC improved property. Producing a plant having an improved property  
CC comprises transforming a plant with a recombinant construct comprising a  
CC promoter region functional in a plant cell operably joined to a  
CC polynucleotide comprising a coding sequence for a polypeptide associated  
CC with the property, and growing the transformed plant. The polypeptide is  
CC useful for improving plant cold tolerance, manipulating growth rate in  
CC plant cells by modification of the cell cycle pathway, improving plant  
CC drought tolerance, providing increased resistance to plant disease,  
CC producing galactomannan (or lignin or plant growth regulators), improving  
CC plant heat tolerance, improving plant tolerance to herbicides, increasing  
CC the rate of homologous recombination in plants, improving plant tolerance  
CC to extreme osmotic conditions or to pathogens or pests, improving yield



CC by modification of photosynthesis, modifying seed oil or protein yield  
CC and/or content, improving yield by modification of carbohydrate, nitrogen  
CC or phosphorus use and/or uptake, or improving yield by providing improved  
CC plant growth and development under at least one stress condition. The  
CC polynucleotide and polypeptide may also be used in recombinant DNA  
CC constructs, in physical arrays of molecules, as plant breeding markers,  
CC or in computer-based storage and analysis systems. The present sequence  
CC is a cotton plant cDNA of the invention. NOTE: The sequence data for this  
CC patent did not form part of the printed specification, but was obtained  
CC in electronic format directly from USPTO at  
CC seqdata.uspto.gov/sequence.html?docID=20040181830. However only 5585  
CC polynucleotide sequences were available, the remaining 52213  
CC polynucleotides and all 58798 protein sequences were not present.  
XX  
SQ Sequence 1211 BP; 366 A; 245 C; 272 G; 328 T; 0 U; 0 Other;

Query Match 38.1%; Score 661; DB 13; Length 1211;  
Best Local Similarity 75.1%; Pred. No. 1.9e-130;  
Matches 851; Conservative 0; Mismatches 280; Indels 2; Gaps 2;

QY 127 ATGAGAGGGAACAAGTTTGGCTGTGATTCGGGTACCTCTCATCTTGGCTGTGTCGCC 186  
DB 80 AATAAAATGGCTAAGCTAAGCTGCGATTCGGGTACCTCTCATCTTGGCTGTGTCGCC 139  
QY 187 TTCACTACACACAGATGCGGCTTTTTCGACACAGTCAGAAATATGAGATGCGCTTGCT 246  
DB 140 TTCAATATACATCCAGATGCGGCTTTTTCACACGCAATCTGAAATATGAGATGCGCA 199  
QY 247 GCTGCAATTTGAAGCAGAAAATCAATTTGACAGCCAGACAGATTCCTTATGACGATT 306  
DB 200 GAAGCGGTGATGAGAGCAGCAATTTGACGAGTCAATGCGACTACTAATTTGATCAAA 259  
QY 307 AGCTGTGACGAGGAAGATAGTTGCTTTCGAAACAAATGAGCGTCAGGACGAG 366  
DB 260 AGTATGCAACAGAGCAAAATCGTGGCCCTGGAAGGGGAGGAGCGAAGGATCAGGAG 319  
QY 367 TGCCGACAAATTAAGGCTTTGTTTTCAGGATCTTGAAGTAAAGGCAATAAAAGTTGATC 426  
DB 320 TGCCGCAATTTGAAGACTTTGTTTAAATGATCTTGAAGAAAGCTCTTCAAGAGTAAT 379  
QY 427 GGAATGTACAGATGCGAGTGGCTGTGTAGTTGTTATGGCTTGCATTCGGGCTGATTAC 486  
DB 380 GACAAACACAGGTACCTGTGCGAGCTGTAGTAATCATGCGGTGCAATCGTGTGATTAT 439  
QY 487 CTGGAAGAAGACTAATAATCCATCTTAAATACCAATATCTTGTGCTGCTCAAAATACCT 546  
DB 440 CTGGAAGGACAGATTTGCTGTCTTAAAGTATCAGAGCTCTGTGCTTTCGAAAGTATCCA 499  
QY 547 CTTTTCATATCCAGGATGATCATCTCTGATGTCAGGAAGCTTGTCTTTCAGCTATGAT 606  
DB 500 CTTTTTGTATCTCAGGATGGATCAGATCAAGGGTTAAACCTAAGGCTTTGAGTTATAG 559  
QY 607 CAGCTGAGCTATATGACGACTTTGATTTTGAACCTGTGCTATCTTGAAGACCCAGGGGAG 666  
DB 560 GAGCTAATTTATGACGACATAGATTATGATCCCGTGCATACAGACCGGCTTGGGAA 619  
QY 667 CTGATTCATATACAAATTCAGCTCATTAAGTGGGCAATTCAGATCAGCTGTTTAC 726  
DB 620 TTGATCGCATACTACAAATTTGCCGCTCACTACAAATGGGCAATTCAGATGAGTTGTTCTAC 679  
QY 727 AAGCATATTTAGCGGTTATCATCTACAGATGATATGGAATTTGCCCTGATTTT 786  
DB 680 AAGCAAAATTTAGCGGAGTAAATACTTGAAGATGATATGGAATTTGCCCTGATTTT 739  
QY 787 TTTGACTTTTTCAGGCTGGAGCTACTCTTTCGACAGACAACTGCAATTTATGCTATT 846  
DB 740 TTTGATTACTTTAGCGAGCTGCTGCCCTTCTCGACAGGACAACTCAATTTATGCTGT 799  
QY 847 TCTTCTGGAATGACAAATGGACAAATTCAGTTTGTCCCAAGATCTTATGCTCTTTACGCG 906  
DB 800 TCTCATGGAATGACAAATGGGCAAAAGCAGTTTGTGTATGACCCATATGCACTTTATGCG 859  
QY 907 TCAGATTTTTCGCCGCTTGGATGGATGCTTTTCAAAATCTACTTTGGGACGAAATTAICT 966

DB 860 TCAGATTTCTTCTGCTGCTTGGCTGGATGCTTTACTATAATCTGTATGGAATGAGCTATCA 919  
QY 967 CCAAACTGCGCAAGGCTTACTGGGACGACTGGCTAAGACTCAAGAGAAATCACAGAGGT 1026  
DB 920 CCAAAATGGCCAAAGCTTACTGGGATGACTGGTGGATTAATAAATAATCACATGGT 979  
QY 1027 CGACAAATTTATTCGCCAGAGTTTCCAGAACATATAATTTTGGTGAGCATGGTTCAGT 1086  
DB 980 CGACAAATTCCTGCTGCTGAAGTATGCAGAACATATAATTTTGGTGAGCATGGTTCAGC 1039  
QY 1087 TTGGGCGAGTTTTCAGCAGATATCTTGAAGCAATTAATTAATGATGTCAGGTTGAT 1146  
DB 1040 ATGGGCGAGTTTTCGAAAAAATACCTTGCACCTATTAAAGATGATGACGT-CAGGTGAC 1098  
QY 1147 TGAAGTCAATGACCTTACTTACCTTTTGGAGGACAAATTACGTGAAACACATTTTGGTGAC 1206  
DB 1099 TGAAGTCCA-GGATTTGAGTACCTAACTGAGGAAAAAATATGCCAATACTTTGCGAGAC 1157  
QY 1207 TTGGTTAAAGGCTAAGCCCATCCATGGAGCTGCTGCTTTGAAAGCATT 1259  
DB 1158 ATTTGAAAGCGCAAAACCTGCTCCCTGGAACAGATTCTGCTCTTATGGCATT 1210

RESULT 8

ADD68118  
ID ADD68118 standard; DNA; 2180 BP.

AC ADD68118;

XX 15-JAN-2004 (first entry)

XX Glycosyltransferase associated hybrid expression cassette SEQ ID NO:35.

XX galactosyltransferase; transmembrane; glycosyltransferase;  
XX glycan processing; ds; hybrid gene expression cassette.

XX Unidentified.

XX W02003078637-A2.

XX 25-SEP-2003.

XX 18-MAR-2003; 2003WO-IB001626.

XX 19-MAR-2002; 2002US-0365735P.

XX (PLAN-) PLANT RES INT BV.

XX Bakker HAC, Florack DEA, Bosch HJ, Rouwendal GJA;

XX WPI; 2003-902819/82.

XX New nucleic acid encoding a hybrid enzyme comprising a transmembrane  
XX region of a plant glycosyltransferase and a catalytic region of a  
XX mammalian glycosyltransferase, useful for optimizing glycan processing in  
XX organism.

XX Disclosure; SEQ ID NO 35; 139pp; English.

XX The invention relates to a novel nucleic acid encoding a hybrid enzyme  
XX comprising a transmembrane region of a plant glycosyltransferase and a  
XX catalytic region of a mammalian glycosyltransferase. The nucleic acid  
XX molecules are useful for optimising glycan processing in an organism so  
XX that a glycoprotein having complex bi-antennary glycans and containing  
XX galactose residues on both arms and which are devoid of xylose and fucose  
XX can be obtained. The present sequence is used in the exemplification of  
XX the invention.

XX Sequence 2180 BP; 648 A; 420 C; 492 G; 620 T; 0 U; 0 Other;

Query Match 37.9%; Score 658.6; DB 10; Length 2180;  
Best Local Similarity 75.9%; Pred. No. 1.4e-189;

Matches		814; Conservative	0; Mismatches	259; Indels	0; Gaps	0;
Qy	409	GGCATAAAAAAGTTGATCGGNAATGTACAGATGCCAGTGCCTGTAGTTGTTATGGCT	468			
Db	882	GGGAAAACTGACTTGGGGACCATGGGCAGATGCGTGTGGCTGTAGTTATGGCC	941			
Qy	469	TGCAATCGGGCTGATTACCTGGAAAAGACTATTAAATCCATCTTAAAAATACCAAAATCT	528			
Db	942	TGCAGTCGTGAGACTATCTTGAAGGACTGTGTAATCAGTTTAAATATCAATCAAACTCCC	1001			
Qy	529	GTTCGCTCAAAATATCCTCTTTTCATATCCAGGATGCATCACAATCCTGATGTCAGGAAG	588			
Db	1002	GTTCGCTCAAAATATCCTCTATTATATCTCAGGATGATCTGATCAAGCTGTCAAGAGC	1061			
Qy	589	CTTCGCTTTGAGCTATGATCAGCTGACGTATATGACGACCTTGGATTTTGAACCTGTGCAT	648			
Db	1062	AAGTCATTTGAGCTATAATCAATTAACATATATGACGACCTTGGATTTTGAACCAAGTGTC	1121			
Qy	649	ACTGAAAGACCCAGGGAGCTGATTGTCATCTACAAAATTCACGTCATTAACAAGTGGGCA	708			
Db	1122	ACTGAAAGGCTGGCGAAGTACTGCTGCTACTACAAGATTGCACTACTCAAGTGGGCA	1181			
Qy	709	TTGGATCAGCTGTTTTACAAGCAATAATTTTAGCCGTGTTATCATCTAGAGATGATATG	768			
Db	1182	CTGGACCAAGTTGTTTTACAACAACAATAATTTAGTCGAGTGATTAATCTAGAAGATGATG	1241			
Qy	769	GAATTCGCCCTGATTTTTTTTGACTTTTTTTGAGGCTGGAGCTACTCTTCTTGACAGAGC	828			
Db	1242	GAATTCGCTCCAGACTTCTTTGATTACTTTTGGGCTGCAGCTAGCTCTCATGGATAGGGAT	1301			
Qy	829	AAGTCGATATTGGCTATTTCTCTTGAATGACAATGACAATAATGCAAGTTTGTCCAAGAT	888			
Db	1302	AAACCAATATTGGCTGCTCATCATGGAATGATTAATGGACAGACGATTTGTGCATGAT	1361			
Qy	889	CCTTATGCTCTTTACCGCTCAGATTTTTTTCCCGGCTCTTGGATGGATGCTTCAAAATCT	948			
Db	1362	CCCTATCGCTATACCGATCAGATTTTTTTTCTGGCCTTGGTGGATGCTCAAGAGATCG	1421			
Qy	949	ACTTGGACGAATATCTCCAAAGTGGCCAAAGCTTACTTGGGACGACTGGCTAAGACTC	1008			
Db	1422	ACTTGGGATGAGTTATCACCAAAGTGGCCAAAGCTTACTTGGGATGATTTGGCTCAGACTA	1481			
Qy	1009	AAAGAGATCACAGAGTGCACAATTTATTCGCCCAAGATTTTGCAGAACATATAATTT	1068			
Db	1482	AAGGAAACCATATAAGGCCGCCAATTCATTCGACCCGGAAGTCTGTAGAACATACAAATTT	1541			
Qy	1069	GGTGAGCATGGTCTAGTTTGGGCGAGTTTTTCAAGCAGTATCTTGAGCCAAATTAACATA	1128			
Db	1542	GGTGAAACATGGTCTAGTTTGGGACAGTTTTTTCAGTCAGTATCTGGAACCTATAAGCTA	1601			
Qy	1129	AATGATGTCAGGTTGATTTGGAAGTCAATGGACCTTTAGTTACCTTTTGGAGGACAATTAC	1188			
Db	1602	AACGATGTACGGTTGACTGGAAAGCAAAAGACCTGGGATACCTGCACAGAGGGAAACTAT	1661			
Qy	1189	GTGAAACACTTTGCTGACTTGTGTTTAAAGGCTAAAGCCCATCCATGGAGCTGATGCTGC	1248			
Db	1662	ACCAAGTACTTTTCTGGCTTAGTGAGACAAGACGACCAATTCAGGTTCTGACCTTGTG	1721			
Qy	1249	TTGAAAGCATTTTAAATAGATGTGATGTGCGTATTTCAGTACAGAGATCAACTAGACTTT	1308			
Db	1722	TTAAAGGCTCAAAACATAAAGGATGATTTTCGATCCGTTATAAGACCAAGTAGATTT	1781			
Qy	1309	GAAATATCGCACCGCAATTTGGCATTTTGAAGATGGAAGATGGTGTACCAAGTGCA	1368			
Db	1782	GAACGCATTGACGGGGAATTTGGTATATTTGAAGATGGAAGGATGGTGTGCCCTCGAACA	1841			
Qy	1369	GCATATAAGGAATAGTAGTTTTCGCTTACCAACGTCACAGCGTGATTTCTTGTGGC	1428			
Db	1842	GCATATAAGAGTAGTGTGTGTTTCGAATCCAGAACAAAGACGTGATTTCTGTGGTGGG	1901			
Qy	1429	CATGATTCGCTTCAACAACTCGGAATTTGAAGATATCTTAAACAAAGATATGATG	1481			
Db	1902	CCAGATTCGTAAATGCAGCTTGGAAATTCGAAATTCCTGATGCGGATCCGCTAG	1954			

RESULT 9		
ADD68117	ADD68117 standard; DNA; 9240 BP.	
XX	AC	ADD68117;
XX	DT	15-JAN-2004 (first entry)
XX	DE	Glycosyltransferase associated triple cassette SEQ ID NO:34.
XX	KW	galactosyltransferase; transmembrane; glycosyltransferase;
XX	KW	glycan processing; ds; triple cassette.
OS		Unidentified.
XX	PN	WO2003078637-A2.
XX	PD	25-SEP-2003.
XX	PF	18-MAR-2003; 2003WO-IB001626.
XX	PR	19-MAR-2002; 2002US-0365735P.
XX	PA	(PLAN-) PLANT RES INT BV.
XX	PI	Bakker HAC, Florack DEA, Bosch HJ, Rouwendal GJA;
XX	DR	WPI; 2003-902819/82.
PT		New nucleic acid encoding a hybrid enzyme comprising a transmembrane
PT		region of a plant glycosyltransferase and a catalytic region of a
PT		mammalian glycosyltransferase, useful for optimizing glycan processing in
PT		organism.
XX		
PS		Disclosure; SEQ ID NO 34; 139pp; English.
XX		
CC		The invention relates to a novel nucleic acid encoding a hybrid enzyme
CC		comprising a transmembrane region of a plant glycosyltransferase and a
CC		catalytic region of a mammalian glycosyltransferase. The nucleic acid
CC		molecules are useful for optimising glycan processing in an organism so
CC		that a glycoprotein having complex bi-antennary glycans and containing
CC		galactose residues on both arms and which are devoid of xylose and fucose
CC		can be obtained. The present sequence is used in the exemplification of
CC		the invention.
XX		
Seq	Sequence 9240 BP; 2707 A; 1778 C; 2037 G; 2718 T; 0 U; 0 Other;	
	Query Match	37.9%; Score 658.6; DB 10; Length 9240;
	Best Local Similarity	75.9%; Pred. No. 3.2e-189;
	Matches 814; Conservative	0; Mismatches 259; Indels 0; Gaps 0;
Qy	409	GGCATAAAAAAGTTGATCGGAATGTACAGATGCCAGTGCCTGCTAGTTGTTATGGCT 468
Db	882	GGGAAAACTGACTTGGGGACCATGGGCAGATGCGCTGTGGCTGTAGTGTATGGCC 941
Qy	469	TGCAATCGGGCTGATTACCTGGAAAAGACTATTAAATCCATCTTAAAAATACCAAAATCT 528
Db	942	TGCAGTCGTGCAGACTATCTTGAAGGACTGTTAAATCAGTTTAAACATATCAAACTCCC 1001
Qy	529	GTTCGCTCAAAATATCCTCTTTTCATATCCAGGATGGATCACAATCCTGATGTCAGGAAG 588
Db	1002	GTTCGCTTCAAAATATCCTCTATTATATCTCAGGATGGSATCTGATCAAGCTGTCAAGAGC 1061
Qy	589	CTTCGCTTTGAGCTATGATCAGCTGACGTATATGCAGCACTTGGATTTTGAACCTGTGCAT 648
Db	1062	AAGTCATTGAGCTATAATCAATTAACATATATGCAGCACTTGGATTTTGACCAAGTGGTC 1121
Qy	649	ACTGAAAGACCCAGGGAGCTGATTGTCATCTACAAAATTCACGTCATTAACAAGTGGGCA 708
Db	1122	ACTGAAAGGCGCTGGCGAAGTCACTGCGGTACTACAAGATTGCGAGTCACTACAAGTGGGCA 1181



Db 852 AAGGAAACCATAAAGCCGCCAATTTCATTCGACCGGAGTCTGTAGAACATACAAATTTT 911  
Qy 1069 GGTGAGCATGTTCTAGTTGGGCGAGTTTTTCAGCAGATATCTTTGAGCCAAATTAACATA 1128  
Db 912 GGTGAACATGGGTCTAGTTTGGGACAGTTTTTCAGTCAGTATCTGGAACCTATAAAGCTA 971  
Qy 1129 AATGATGTCGACGGTGTGATTCGAGTCAATGACCTTAGTTACCTTTTGGAGGACAATTAC 1188  
Db 972 AACGATGTGACGGTGTGATTCGAGTCAATGACCTTAGTTACCTTTTGGAGGACAATTAC 1031  
Qy 1189 GTGAAACACATTTGGTACTTTGGTGTAAAGGCTAAGCCCATCCATCGAGCTGATGCTGTC 1248  
Db 1032 ACCAAGTACTTTTCTGGCTTAGTGAGACAAGCAGCAACCAATTCAAGTCTTGACCTTGT 1091  
Qy 1249 TTGAAAGCATTTATACATAGATGTTGATGTCGTTATTCAGTACAGAGATCAACTAGACTTT 1308  
Db 1092 TTAAGGCTCAAAACATATAAGGATGATGTTCTGATCCGGTATAAAGACCAAGTAGAGTTT 1151  
Qy 1309 GAAATATCCGACCGCAATTTGGCATTTTTCAGAAATGGAAGGATGGTACCACCGTGA 1368  
Db 1152 GAACGCAATTCAGGGGAATTTGGTATATTTGAAGATGGAAGGATGGTGGCTCGAACA 1211  
Qy 1369 GCATATAAGGAATAGTAGTTTTCGGTACCAACGTCAGACGCTGATTTCTTGTGGC 1428  
Db 1212 GCATATAAGGATAGTGTGTTTTCGAATCCAGACACAGACGCTGATTTCTGTTGG 1271  
Qy 1429 CATGATTGCTTCAACAACTCGGAATTTGAAGATATCTTAAACAAAGAT 1474  
Db 1272 CCAGATTCTGTAATGCAGCTTGGAAATTCGAAATTCCTGATCGCGAT 1317

RESULT 11

ADD68136  
ID ADD68136 standard; DNA; 1119 BP.  
XX AC ADD68136;  
XX DT 15-JAN-2004 (first entry)  
XX DE Plant GntI/mouse signal peptide hybrid gene SEQ ID NO:53.  
XX KW galactosyltransferase; transmembrane; glycosyltransferase;  
XX KW glycan processing; ds; gene; plant; mouse.  
XX OS Chimeric.  
XX OS Arabidopsis thaliana.  
XX OS Mus sp.  
XX PH Key Location/Qualifiers  
XX FT CDS 1..1119  
XX FT /\*tag= a  
XX FN WO2003078637-A2.  
XX PD 25-SEP-2003.  
XX PF 18-MAR-2003; 2003WO-IB001626.  
XX PR 19-MAR-2002; 2002US-0365735P.  
XX PA (PLAN-) PLANT RES INT BV.  
XX PI Bakker HAC, Florack DEA, Bosch HJ, Rouwendal GJA;  
XX DR WPI; 2003-902819/82.  
XX DR P-PSDB; ADD68136.  
XX XX  
XX PT New nucleic acid encoding a hybrid enzyme comprising a transmembrane  
XX PT region of a plant glycosyltransferase and a catalytic region of a  
XX PT mammalian glycosyltransferase, useful for optimizing glycan processing in  
XX PT organism.  
XX PS Disclosure; SEQ ID NO 53; 139pp; English.

XX The invention relates to a novel nucleic acid encoding a hybrid enzyme  
CC comprising a transmembrane region of a plant glycosyltransferase and a  
CC catalytic region of a mammalian glycosyltransferase. The nucleic acid  
CC molecules are useful for optimising glycan processing in an organism so  
CC that a glycoprotein having complex bi-antennary glycans and containing  
CC galactose residues on both arms and which are devoid of xylose and fucose  
CC can be obtained. The present sequence is used in the exemplification of  
CC the invention.  
XX  
Qy Sequence 1119 BP; 326 A; 203 C; 278 G; 312 T; 0 U; 0 Other;  
Query Match 37.8%; Score 656.6; DB 10; Length 1119;  
Best Local Similarity 77.0%; Pred. No. 4e-189;  
Matches 800; Conservative 0; Mismatches 239; Indels 0; Gaps 0;  
Qy 425 TCAGAAATGTACAGATGCCAGTGGCTGTAGTGTCTTATGGCTTGCAATCGGGCTGATT 484  
Db 65 TCGACATGGGACAGATGCCGTGGCTGTAGTGTCTTATGGCTGCAGTCGTGCAGACT 124  
Qy 485 ACCTGAAAGACTATTAATCCATCTTAAATACCAATATATCTGTTGCGTCAAAATATC 544  
Db 125 ATCTTGAAGGACTGTAAATCAGTTTAAACATATCAAACTCCCGTTGCTTCAAAATATC 184  
Qy 545 CTCTTTTCATATCCAGGATGGATCACATCCTGATGTCAGGAAGCTTGTCTTGAGCTATG 604  
Db 185 CTCTATTATATCTCAGGATGGATCTGATCAAGCTGTCAAGAGCAAGTCAATTGAGCTATA 244  
Qy 605 ATCAGCTGACGTATATGCAGCACTTGGATTTTGAAGCTGTGCATATCTGAAAGACCGAGG 664  
Db 245 ATCAATTAACATATATGCAGCACTTGGATTTTGAACAGTGGTCACTGAAAGCGCTGGCG 304  
Qy 665 AGCTGATGCTACTACAAATTTGACGCTCATTTAAAGTGGGCAATTTGGATCAGCTGTTTT 724  
Db 305 AACTGCTCGTACTACAAAGATTGACGCTCACTACAAAGTGGGCACTGGACAGTTGTTTT 364  
Qy 725 ACAAGCATATTTTAGCCGTGTATCATATAGAGATGATATGGAATTTGCCCTGATT 784  
Db 365 ACAACACAAATTTAGTCAGTGATTATATAGAGATGATATGGAATTTGCTCCAGACT 424  
Qy 785 TTTTGTGACTTTTGTAGGCTGGAGCTACTCTTCTTGACAGAGACAAAGTCCGATTATGGCTA 844  
Db 425 TCTTGTATTACTTTGAGGCTGCAGCTAGTCTCATGATAGGATATAAACCATTTATGGCTG 484  
Qy 845 TTTCTTCTTGGAAATGACAAATGACAAATGACGTTTGTCCAAAGTCTTATGCTCTTTACC 904  
Db 485 CTTTCATCGAATGATTAATGGACAAAGCAGTTTGTGCATGATCCCTATGCGCTATACC 544  
Qy 905 GCTCAGATTTTTCCTCCGGTCTTGGATGGATGCTTTTCAAAATCTACTTGGGACGAAATTAT 964  
Db 545 GATCAGATTTTTCCTCCGGTCTTGGATGGATGCTTTCAAAATCTACTTGGGACGAAATTAT 604  
Qy 965 CTCGAAAGTGGCCAAAGGCTTACTGGGACGACTGGCTTAAGACTCAAAGAGAAATCACAGAG 1024  
Db 605 CACCAAGTGGCCAAAGGCTTACTGGGATGATTTGGCTGAGACTTAAAGGAAACCATAAAG 664  
Qy 1025 GTCGACAAATTTATTCGCCAGAAAGTTTGCAGAAACATATATATTTTGGTACGATGTTCTA 1084  
Db 665 GCGCCAAATTCATTTCGACCGGAAGTCTGTAGAACATACAAATTTTGGTGAACATGGGTCTA 724  
Qy 1085 GTTTGGGCGAGTTTTTCAAGCAGTATCTTGAGCCCAATTAAGCTTAAATGATGTCCAGGTG 1144  
Db 725 GTTTGGGACAGTTTTTTCAGTCAGTATCTGGAACCTATAAGCTTAAACGATGTGACGGTTG 784  
Qy 1145 ATTGGAAGTCAATGGACCTTAGTTTACCTTTTGGAGGACAAATTAAGCTTAAACGATGTG 1204  
Db 785 ACTGGAAGCAAGGACCTGGGATACCTGCAGAGGGGAACTATACCAAGTACTTTTCTG 844  
Qy 1205 ACTTGTTTAAAGGCTTAAGCCCATCCATGGAGCTGATGCTGTCTGAAAGCATTTAAACA 1264  
Db 845 GCTTAGTGAGACAAAGCAGCAACCAATTCGAGGTTCTGACCTTGTCTTAAAGGCTCAAAACA 904  
Qy 1265 TAGATGATGTGCGTATTTCAGTACAGAGATCAACTAGACTTTTGAANAATATCGCAGCGC 1324

Accession	Gene	Strain	Position	Sequence
905	Db	TAAAGGATGATGTTCTCGTATCCGGTATAAAGACCAAGTAGAGTTTGAACGCAATTGCAGGGG	964	
1325	Qy	AATTTGGCATTTTGAAGAATGGNAGATGGTGTACACGTGCGAGCATATTAAGGAATAG	1384	
965	Db	AATTTGGTATATTTGAAGAATGGAAGATGGTGTGCCCTCGAACAGCATATAAAGAGGATAG	1024	
1385	Qy	TAGTTTTCCGGTACCAAACGTCAGACGTGTATTCCTTGTGTGGCCATGATTCGGTTCAAC	1444	
1025	Db	TGGTGTTCGAATCCAGACAAACAGACGTGTATTCCTGTTGGGCCAGATTCCTGTAATGC	1084	
1445	Qy	AATCGGAATTGAAGATAC	1463	
1085	Db	AGCTTGAATTCGAAATTC	1103	

RESULT 12  
ADT20185/C  
ID ADT20185 standard: CDNA: 1898 BP.

AA ADT20185;  
AC

13-JAN-2005 (first entry)

XX DE Plant cDNA, Seq ID 5511.

xx Plant; ss: gene; transgenic; cold tolerance; growth rate;  
 kw drought tolerance; disease resistance; galactomannan production;  
 kw plant growth regulator; heat tolerance; herbicide tolerance;  
 kw lignin production; extreme osmotic condition tolerance;  
 kw pathogens resistance; pest resistance; yield improvement; seed oil yield;  
 kw seed protein yield.

XX QS Viridiplantae.

AA  
PN US2004216190-A1.XX  
PD  
28-OCT-2004.

18-DEC-2003; 2003US-00739930.

XX  
PR 28-APR-2003: 2003US-00424599.

PR 28-APR-2003; 2003US-00425115.

PA (KOVA/) KOVALIC D K.

PI Kovalic DK;

WPI; 2004-757369/74.

New recombinant DNA constructs useful in the field of biochemistry and genetics, and in particular for producing transgenic plants with improved biological characteristics.

PS Claim 1: SEO ID NO 5511: 14pp; English.

The invention relates a recombinant DNA construct comprising a polynucleotide having any of 5544 nucleotide sequences (cDNAs SEQ ID NO: 1-5544) and encoding a polypeptide with any of 5544 amino acid sequences (SEQ ID NO: 5545-11088). The cDNAs and proteins are from corn, soybean, Arabidopsis, wheat and rape but the specification does not indicate which sequences is derived from which organism. Also included is a method of producing a plant having an improved property, comprising transforming a plant with a recombinant DNA construct comprising a promoter region functional in a plant cell operably joined to a polynucleotide encoding a polypeptide associated with the property, and growing the transformed plant. The property is selected from improving plant cold tolerance, for manipulating growth rate in plant cells by modification of the cell cycle pathway, for improving plant drought tolerance, for providing increased resistance to plant disease, for galactomanan production, for production of plant growth regulators, for improving plant heat tolerance, for improving plant tolerance to herbicides, for increasing the rate of homologous recombination in plants, for lignin production, for improving

Db 1118 ACTACTTTGAGCTGCAGCGAAATTTACTTGACACTGACAGACAATAATGGCTGTTTCAT 1059  
 Qy 851 CTTGGAATGCAATGGGCAAAATGAGTTTGTCCAGATCTTATGCTCTTTACCGCTCAG 910  
 Db 1058 CTTGGAATGCAATGGGCAAAAGCAGTTGCTTTATGACCCCAAGGCTCTTTACCGTTCCG 999  
 Qy 911 ATTTTTCCTCCGGTCTTTGGATGGATGCTTTTCAAAATCTACTTGGGACGAATTTCTCCAA 970  
 Db 998 ACTTCTTTCCGGGCTTTGGATGGATGCTTAAGAAAGTCAACATGATGAGACTGTCGCCAA 939  
 Qy 971 AGTGCCCAAAAGGCTTACTGGGACGACTGGCTAAGACTCAAGAGAAATCACAGAGGTCGAC 1030  
 Db 938 AGTGCCCAAAAGCTTATTGGGATGACTGGGTGAGGCTAAAGAGGTACACAGAGATCGC 879  
 Qy 1031 AATTATTCCGCCAAGTTTGCAGAACATATAATTTTGGTGAGCATGGTCTAGTTTG 1090  
 Db 878 AGTTTATTCCGCCAAGATATGCGAGGACATACAACTTTGGCGAGCATGGATCAAGCATGG 819  
 Qy 1091 GGCAATTTTCAAGCAGTATCTTGAGCCAAATTAACACTTAAATGATGTCAGGTTGATTGGA 1150  
 Db 818 GACAATCTTTGATCAATATCTTGAACCTATCAAGTTAAATGATGCTCATATTGACTGGA 759  
 Qy 1151 AGTCAATGGACCTTAGTTACCTTTTGGAGGACAAATACGTGAACACTTTTGGTACTTGG 1210  
 Db 758 ATTCGAGGACCTGAGCTACCTCAAGGAGGACAAAGTTTGTGACCAATTCGGGAAAGACG 699  
 Qy 1211 TTAAGAGGCTAAGCCATCATGAGGCTGATGCTGCTTGAAGAGATTTAAACATAGATG 1270  
 Db 698 TGGCTAGCGCCACACCTGTGCTGATGGATCCGATGCTCTGTTGAAGGCCCAACAATCTGGATG 639  
 Qy 1271 GTGATGTGCGTATTTCAGTACAGAGATCAACTAGACTTTTGAATAATATCGACGCGCAATTTC 1330  
 Db 638 TGGACGTAAGATTCAGTATGACAAATCAGGCGGACTTTTGAGCGTATAGCTCGGCAGTTTG 579  
 Qy 1331 GCATTTTGAAGATGGAAGGATGGTGTACACGCTGCAGCATATAAAGGAATAGTAGTTT 1390  
 Db 578 GAATATTGGAAGAGTGAAGGACGGTGTTCACGGCGGCTTACAAAGCGGTGTTGGTGT 519  
 Qy 1391 TCCGGTACCA---AACGTCGACGCTGATTTCTTTGTTGGCATGATTCGCTTCAACAAC 1447  
 Db 518 TCCGGTCAAGGGCAGCGGAGACGGATATACCTGTGGTCCGCGTCTCTTCGCCAAC 459  
 Qy 1448 TCGGAATTGGAAGATA 1462  
 Db 458 TCGGGTTAGCTTA 444

RESULT 13  
 ID AEB65577 standard; DNA; 896 BP.  
 AC AEB65577;  
 XX  
 DT 22-SEP-2005 (first entry)  
 XX  
 DE Rice genome derived DNA sequence, SEQ ID 722.  
 XX  
 KW transcription; gene regulation; transgenic plant; RNA interference;  
 XX transformation; antibody; ds.  
 OS Oryza sp.  
 XX  
 PN JP2005185101-A.  
 XX  
 PD 14-JUL-2005.  
 XX  
 PF 11-DEC-2002; 2002JP-00383870.  
 XX  
 PR 30-MAY-2002; 2002JP-00203269.  
 XX  
 PA (DOKU-) DOKURITSU GYOSHI HOJIN NOGYO SEIBUTSU SH.  
 PA (SEIB-) SEIBUTSUKI TOKUTEI SANGYO GIJUTSU.  
 PA (DOKU-) DOKURITSU GYOSHI HOJIN RIKAGAKU KENKYUSH.

PA (KOKU-) ZH KOKUSAI KAGAKU SHINKO ZAIDAN.  
 XX  
 PI Kikuchi H, Hayashizaki Y, Otono Y, Matsuura K, Murakami K;  
 PI Kishimoto N, Sato K, Nagata T, Kawakami N, Yazaki J, Ishikawa M;  
 PI Doi K, Kawai J;  
 XX  
 DR WPI; 2005-566181/58.  
 XX  
 PT Novel DNA encoding transcription factor, derived from rice plant, useful  
 PT for obtaining transcriptional-regulatory regions in plant and for  
 PT producing modified plant.  
 XX  
 PS Claim 1; SEQ ID NO 722; 2928bp; Japanese.  
 XX  
 CC The invention relates to a novel DNA sequence encoding a transcription  
 CC factor derived from a plant. The invention further comprises antisense  
 CC RNA sequences, ribozyme activity RNA, RNAi sequences, a vector,  
 CC transformed plant cells, antibodies and proteins, all related to the  
 CC novel plant DNA sequences of the invention. The novel DNA is preferably  
 CC derived from a rice-genome database. The invention further provides a  
 CC method for determining the transcription regulatory regions of the rice  
 CC genome. The novel DNA is useful for controlling the expression of a gene  
 CC in a plant and for producing a modified plant with desired and different  
 CC characteristics. The plant DNA and method enables the acquisition of many  
 CC transcriptional-regulatory regions. This polynucleotide represents a DNA  
 CC sequence taken from a rice genome clone library for use in the invention.  
 CC Note: This sequence is not shown in the specification. It has been  
 CC retrieved from a sequence listing in electronic format from the Japanese  
 CC Patent Office. The invention claims DNA sequences of SEQ ID Nos 1 to  
 CC 28469 and encoded protein sequences of SEQ ID Nos 28470 to 56791,  
 CC however, the sequence listing only provided the DNA sequences of SEQ ID  
 CC Nos 1 to 3032.  
 XX  
 SQ Sequence 896 BP; 256 A; 184 C; 220 G; 236 T; 0 U; 0 Other;  
 Query Match 24.5%; Score 425.8; DB 14; Length 896;  
 Best Local Similarity 73.6%; Pred. No. 1e-118;  
 Matches 557; Conservative 0; Mismatches 197; Indels 3; Gaps 1;  
 Qy 695 ATTCAAGTGGGCAATGGATGAGCTGTTTACAGCAATATTTAGCGTGTATCATAC 754  
 Db 1 ATTCAAAATGGGCTTTGGATGAGCTTATCATCAACACAACTTGGCGGAGTAATAATTC 60  
 Qy 755 TAGAAGATGATGGAATGGCCCTGATTTTTCACATTTTGGAGTGGGCTACTC 814  
 Db 61 TGAAGATGATGGAATGGGCTGATTTTTCACATTTTGGAGTGGGCTACTC 120  
 Qy 815 TTCTTGACAGACAAGTGGATGATGCTTATGCTTCTTGGATGACAATGGAACAATGC 874  
 Db 121 TACTTGATTAACGATAGACAATCATGCGCGTCTTCTTGGATGACAATGGAACAAGC 180  
 Qy 875 AGTTTGTCCAAGATCCTTATGCTTTTACCGTCAAGATTTTTCGGTCTTTGGATGGA 934  
 Db 181 AATTTGTTTACGACCCCAAGGCTCTTTACCGTGGGATTTCTTTCTCGGACTTGGATGA 240  
 Qy 935 TGCTTTCAAAATCTACTTGGGAGCAATATCTCAAGTGGCCCAAGGCTTACTGGGACG 994  
 Db 241 TGTTAAAGAGGCTTACATGATTTGAGCTGTACCTAAGTGGCCCAAGCTTATTTGGATG 300  
 Qy 995 ACTGGCTAAGACTCAAGAGAAATCACAGAGGTCGACAATTTATTCGCCCAAGAGTTTGA 1054  
 Db 301 ACTGGGTGAGGCTTAAAGAGGTTACAGAGACCGGCAATTTATTCGCCCGGAGTCTGA 360  
 Qy 1055 GAACATATAATTTTGGTGAGCATGTTCTAGTTTGGGCGAGTTTTCAGACAGTATCTTG 1114  
 Db 361 GAACATATAATTTTGGGCAACATGTTGTTCAAGCATGCGGCAATTTCTTTCAGACAGTCTGG 420  
 Qy 1115 AGCAATTAACCTAAGTATGTCAGGTTGATTTGGAAGTCAATGGACCTTAGTACTCTT 1174  
 Db 421 AACCAATCAAGCTTAAATGATGCTCCATATCAAGTGGAAATTTCTGAGGACCTGAGCTTCA 480  
 Qy 1175 TGAGGACAAATTTACGTGAACACACTTTGGTGTGACTTGGTTAAAGAGGCTTAAGCCCATCATG 1234

Db 481 AGGAGCAAGTCTCTGATCCAAATTTGGGAAGACGTCGCTAGTGCCACCCCTCTCCATG 540

Qy 1235 GAGCTGATGCTGCTTGAAGCAATTTAAACATAGATGGTATGCGGTATTCAGTACAGAG 1294

Db 541 GATCTGATGCGCGTGTGAAGCCACACATATGGAATGGATGCGGATGTAAAGGATCCAGTATAACG 600

Qy 1295 ATCAACTAGACTTTGMAAATATCGCACGGCAATTTGGCATTTTGAAGATGGAAGGATG 1354

Db 601 ATCAGGAAGACTTTGAGCGGATAGCTCTCAATTTGGAATATTTGAAGAATGGAAGGATG 660

Qy 1355 GTGTACCACTGTCAGCATATATAAGAAATAGTAGTTTTCGGTACCA---AACGTCACAGAC 1411

Db 661 GCATTCGAAGAACAGCTTATAAAGGAGTAGTAGTCTTCGGGTACAAGAGTAGTGAAGAC 720

Qy 1412 GTGTATTCCTTTGTTGGCCATGATTCGCTTCAACAACT 1448

Db 721 GAATATACCTCTGTTGGTCCGGATTCCCTCAGTCAGCT 757

RESULT 14

AAx78003

ID AAX78003 standard; cDNA; 510 BP.

XX AAX78003;

AC

XX

DT 19-AUG-1999 (first entry)

XX

DE A. thaliana GnTI-homologue cDNA fragment.

XX

XX GnTI; beta-1,2-N-acetylglucosaminyltransferase; glycoprotein; plant;

XX deficient; defective; detection; transgenic plant; sugar residue;

XX medicine; antigenicity; deglycosylation; ds.

XX Arabidopsis thaliana.

OS

XX

PN DE19754622-A1.

XX

PD 10-JUN-1999.

XX

PF 09-DEC-1997; 97DE-01054622.

XX

PR 09-DEC-1997; 97DE-01054622.

XX

PA (VSCH/) VON SCHAEWEN A.

XX

PI Von Schaewen A;

XX

XX WPI; 1999-338905/29.

DR P-PSDB; AAY08890.

XX

XX Nucleic acid encoding plant N-acetylglucosaminyl transferase I useful for

PT generating plants producing glycosylated proteins of low antigenicity.

PS

XX Claim 5; Page 27; 37pp; German.

XX This invention describes novel plant N-acetylglucosaminyl transferase I

CC (GnTI) proteins and their encoding nucleic acids. The nucleic acid

CC sequences of the invention may be used for recombinant production of the

CC encoded proteins, which are then used to raise antibodies (Ab) for

CC identifying plants with deficient or defective production of the protein.

CC They may also be used to detect such plants by hybridization and to

CC isolate related sequences from other plants or to generate antisense or

CC sense constructs for reducing/deleting GnTI protein activity in plants.

CC These transgenic plants may be used to produce glycoproteins with

CC minimal, uniform and defined sugar residues. Such glycoproteins are

CC useful in medicine and research, e.g. human glucocerebrosidase for

CC treating Gaucher's disease. Plants which are defective or deficient in

CC production of the GnTI protein can be made to produce glycoproteins with

CC minimal, uniform and defined sugar residues, of low antigenicity. Use of

CC these plants eliminates the need for the difficult isolation and

CC deglycosylation of native proteins or preparation in defective animal

CC cells. This sequence represents cDNA which encodes an Arabidopsis

CC thaliana GnTI homologue

XX SQ Sequence 510 BP; 143 A; 104 C; 114 G; 149 T; 0 U; 0 Other;

Query Match 19.2%; Score 334; DB 2; Length 510;

Best Local Similarity 78.4%; Pred. No. 8e-91; Indels 0; Gaps 0;

Matches 400; Conservative 0; Mismatches 110;

Qy 424 ATCGAAATGTACAGATGCCAGTGGCTGCTGTAGTTGTATGGCTTGCATTCGGGCTGAT 483

Db 1 ATCGAAAGCTTGGATCCCGAGTGGCGCTGTAGTTGTATGGCTTGCATTCGGGCTGAT 60

Qy 484 TACCTGAAAAGACTATTAATCCATCTTAAATACCAATATCTGTGCGTCAAAATAT 543

Db 61 TATCTTGAAGAGACTGTAAATCACTGTAAATCACTCAATCACTCCCGTTCCTTCAAAATAT 120

Qy 544 CCTCTTTTCAATCCAGGATGGATCACATCCTCATGTGCAGAAAGCTTGTCTTTCAGCTAT 603

Db 121 CCTCTATTTATCTCAGGATGGATCTGATCAAGCTGTCAAGAGCAAGTCAATTCAGCTAT 180

Qy 604 GATCAGCTGAGTATATGCAGCACTTGGATTTTGAACCTGTGCATATCTGAAAGACCGG 663

Db 181 AATCAATTAACATATATGCAGCACTTGGATTTTGAACCGAGTGGTCACTGAAAGCGCTGGC 240

Qy 664 GAGCTGATTCATATACAAAATTTGCAGCTCATTTACAAGTGGGCAATTCGATCAGCTGTTT 723

Db 241 GAACTGACTGCTACTACAAGATTGCAGCTCCTACAAAGTGGGCACTGCAGCCAGTTGTTT 300

Qy 724 TACAAGCATATATTTAGCGGTGTATCATCTAGAAGATGATATGAAATATGCCCTCAT 783

Db 301 TACAAACACAAATTTAGTCGAGTATATATAGAGATGATATGGAATATGCTCCAGAC 360

Qy 784 TTTTTCGACTTTTTCGAGCTGGAGCTACTCTTCTTGACAGAGCAAGTCGATATGCT 843

Db 361 TCTTTGATTACTTTGAGCTGAGCTAGTCTCATGGATAGGATAAAACCATTTATGGCT 420

Qy 844 ATTCTCTTTCGAAATGACAAATGGACAAATGCAGTTTGTCCAAGATCCTTATGCTCTTAC 903

Db 421 GCTTCATCATGGACTGATAATGGACAGAGCAGTTTGTGCATGATTCCTATGCGCTATAC 480

Qy 904 CGCTCAGATTTTTCGCGTCTTGGATGG 933

Db 481 CGATCAGATTTCTTCCCTGGCCACGGCTGG 510

RESULT 15

ADS70553

ID ADS70553 standard; cDNA; 278 BP.

XX AC

XX ADS70553;

XX

DT 18-NOV-2004 (first entry)

XX

XX Corn seedling-derived polynucleotide (cpds), SEQ ID 5569.

DE

XX Corn; seedling-derived polynucleotide; cdp; ss; plant; seed growth;

XX seed development; disease resistance; insect infestation; fungal disease;

XX bacterial infection; Goss' Bacterial wilt; blight;

XX Stewart's bacterial wilt; Holcus spot; bacterial leaf blight; leaf spot;

XX Bacterial stripe; maize dwarf mosaic virus infection;

XX environmental stress; water stress; pH stress; temperature stress;

XX pollution; injury; pesticide.

OS Zea mays.

XX

XX US2003237110-A9.

PN

XX

PD 25-DEC-2003.

XX

XX 06-AUG-2001; 2001US-00923876.

PF

XX

XX 12-MAY-1998; 98US-0085331P.

PR

XX 21-APR-1999; 99US-00298329.

XX

PA (INCY-) INCYTE PHARM INC.  
XX Lalgudi RV, Ito LY, Sherman BK;  
XX WPI; 2002-195165/25.  
XX  
XX New corn seedling-derived polynucleotides and polypeptides, useful in  
PT identifying and altering desired characteristics associated with growth  
PT and development, disease resistance, environmental adaptability, quality  
PT and yield.  
XX  
XX Claim 1; SEQ ID NO 5569; 33pp; English.  
XX  
XX The invention relates to a corn seedling-derived polynucleotide (cdp)  
CC selected from ADS64985-ADS71316, or their complements and fragments. Also  
CC included are a composition for the detection of altered expression of a  
CC cdp (comprising a polynucleotide selected from ADS64985-ADS71316), a  
CC method of detecting a polynucleotide in a biological sample using a cdp,  
CC a method for using oligomers (and amplification) to recover a regulatory  
CC element from a DNA library using oligomers designed against a cdp, a  
CC seedling specific regulatory element that regulates the expression of a  
CC cdp, an expression vector containing a cdp or regulatory element, a plant  
CC transformed with the vector, a host cell containing the vector (and  
CC expressing a corn seedling derived protein, CDP), an anti-CDP antibody,  
CC identifying a compound which binds a CDP and screening a plurality of  
CC compounds for binding to cdp polynucleotide. The cdp polynucleotides,  
CC proteins, vectors, cells and antibodies are useful for the  
CC identification, evaluation and alteration of seed growth and development,  
CC disease resistance (e.g. to insect infestation, fungal disease, bacterial  
CC infection, Goss' Bacterial Wilt, blight, Stewart's bacterial wilt, Holcous  
CC spot, bacterial leaf blight, leaf spot, bacterial stripe and maize dwarf  
CC mosaic virus infection) and resistance to environmental stress (e.g.  
CC water stress, pH stress, temperature stress, pollution, injury or  
CC pesticides. The present sequence is cdp cDNA sequence.  
XX  
SQ Sequence 278 BP; 87 A; 52 C; 60 G; 79 T; 0 U; 0 Other;  
  
Query Match 10.5%; Score 182; DB 7; Length 278;  
Best Local Similarity 78.4%; Pred. No. 1.4e-44;  
Matches 218; Conservative 0; Mismatches 60; Indels 0; Gaps 0;  
  
Qy 625 CAC TTGGATTTTGAACCTGTGCATCTGAAGACCGAGGAGCTGATTGCATCTACAAA 684  
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
1 CATGTGGATCTTGAACCTGTGCAAACTGAAGAGCGGAGGAGATTGACGATATTACAAG 60  
  
Qy 685 ATTGCAGCTCATTACAAGTGGGCATTTGGATCAGCTGTGTTTACAAGCATAAATTTAGCCGT 744  
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
61 ATTGCTAAACATTATTAAGTGGCCTTGACAACTATTTCATTAACATATTTTGCTCGA 120  
  
Qy 745 GTTATCATCTAGAAGATGATATGGAATTCGCCCTGATTTTGTGACTTTTGTAGGCT 804  
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
121 GTAATCATCTTGAAGATGACATGGAGATTGCCCGGATTTTTCGACTACTTCGAGGCT 180  
  
Qy 805 GGAGCTACTCTTCTTGACAGACAAGTCGATTATGGCTATTCTCTTGAATGACAAT 864  
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
181 GCAGCTAACTTCTTGATATGATGACAGACGATTATGGCTGTTCTCTTGAATGACAAT 240  
  
Qy 865 GGACAAATGCAAGTTGTCCCAAGATCCTTATGCTCTTTA 902  
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
241 GGGCAGAAGCAGTTTGTTAACGACCCCAAAAGCTCTTTA 278  
  
Search completed: December 14, 2005, 08:52:03  
Job time : 1011 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 14, 2005, 08:35:23 ; Search time 1307 Seconds  
(without alignments)  
10989.984 Million cell updates/sec

Title: US-10-615-144-3  
Perfect score: 1737  
Sequence: 1 gaattcgccgcgcattga.....ataatcgccgcgcgaattc 1737

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 9793542 seqs, 4134689005 residues

Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications NA Main:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1737	100.0	1737	8	US-10-615-144-3
2	1255	72.3	1669	8	US-10-615-144-1
3	767.2	44.2	1854	7	US-10-615-144-5
4	714.2	41.1	1836	7	US-10-425-114-29332
5	668.2	38.5	1594	7	US-10-437-963-21004
6	661.8	38.1	1862	7	US-10-425-114-487
7	661.8	38.1	2034	8	US-10-425-115-25820
8	661.8	38.1	1211	8	US-10-767-795-2499
9	614.6	35.4	2293	7	US-10-424-599-93080
10	581.2	33.5	1898	8	US-10-739-930-5511
11	577	33.2	1414	7	US-10-424-599-93082
12	527.8	30.4	954	7	US-10-424-599-93078
13	343.6	12.8	1849	7	US-10-437-963-21003
14	222	12.8	769	7	US-10-767-701-5996
15	182	10.5	278	3	US-09-923-876-5569
16	182	10.5	278	3	US-09-923-876-5569
17	181.8	10.5	2580	5	US-10-087-192-374
18	181.8	10.5	31477	5	US-10-087-192-373
19	146.4	8.4	2670	9	US-10-185-475-34
20	146.4	8.4	2670	9	US-10-713-970-12
21	146.4	8.4	2753	8	US-10-087-192-377
22	146.4	8.4	2937	8	US-10-844-874-13
23	146.4	8.4	2937	9	US-10-887-553A-329

24	146.4	8.4	4677	7	US-10-185-475-36
25	146.4	8.4	4677	9	US-10-713-970-14
26	146.4	8.4	32323	5	US-10-087-192-376
27	136.6	7.9	2367	10	US-11-097-143-31280
28	122.6	7.1	590	7	US-10-021-323-15672
29	122.6	7.1	590	8	US-10-767-795-2500
30	105.8	6.1	509	7	US-10-767-701-21087
31	69.8	4.0	347	3	US-09-960-352-2775
32	58.8	3.4	4721	10	US-11-097-143-31279
33	58.6	3.4	613	8	US-10-363-345A-30605
34	58.6	3.4	613	8	US-10-363-345A-30606
35	58.6	3.4	613	9	US-10-363-483A-30605
36	58.6	3.4	613	9	US-10-363-483A-30606
37	55.2	3.2	587	8	US-10-357-930-45887
38	55	3.2	403	8	US-10-357-930-16061
39	39.6	2.3	613	8	US-10-363-345A-30607
40	39.6	2.3	613	8	US-10-363-345A-30608
41	39.6	2.3	613	9	US-10-363-483A-30607
42	39.6	2.3	613	9	US-10-363-483A-30608
43	38.8	2.2	535	5	US-10-027-632-91201
44	38.8	2.2	535	5	US-10-027-632-304430
45	38.8	2.2	535	6	US-10-027-632-91201

ALIGNMENTS

RESULT 1  
US-10-615-144-3  
Sequence 3, Application US/10615144  
Publication No. US20040181927A1  
GENERAL INFORMATION:  
APPLICANT: Von Schaeuwen, Antje  
TITLE OF INVENTION: Plant GntI Sequences and the Use Thereof for the Production of Plants Having Reduced or Lacking N-acetyl Glucosaminyl  
TITLE OF INVENTION: of Plants Having Reduced or Lacking N-acetyl Glucosaminyl  
FILE REFERENCE: 032266-003  
CURRENT APPLICATION NUMBER: US/10/615,144  
CURRENT FILING DATE: 2003-07-09  
PRIOR APPLICATION NUMBER: US/09/591,466  
PRIOR FILING DATE: 2000-06-09  
PRIOR APPLICATION NUMBER: EP 98/08001  
PRIOR FILING DATE: 1998-09-12  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 3  
LENGTH: 1737  
TYPE: DNA  
ORGANISM: Nicotiana tabacum  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (733)...(741)  
OTHER INFORMATION: function: Asn codon in this context is a potential  
OTHER INFORMATION: glycosylation site;  
OTHER INFORMATION: product: N-glycosylation consensus sequence;  
OTHER INFORMATION: phenotype: N-glycans modulate protein properties;  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (733)...(741)  
OTHER INFORMATION: standard name: N-glycosylation site;  
OTHER INFORMATION: label: pot-CHO;  
OTHER INFORMATION: note: GntI sequences from animals do not contain  
OTHER INFORMATION: this feature.  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (127)...(1467)  
OTHER INFORMATION: codon start: 127;  
OTHER INFORMATION: function: initiates complex N-glycans on secretory  
OTHER INFORMATION: glycoproteins;  
OTHER INFORMATION: EC\_number: 2.4.1.101;  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (127)...(1467)

= instant gghh

```
; OTHER INFORMATION: product: beta-1,2-N-acetylglucosaminyltransferase I;  
; OTHER INFORMATION: evidence: EXPERIMENTAL;  
; OTHER INFORMATION: gene: cgi;  
; OTHER INFORMATION: standard_name: gntI;  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (127)...(1467)  
; OTHER INFORMATION: label: ORF;  
; OTHER INFORMATION: note: first gntI sequence from tobacco (unpublished).  
; FEATURE:  
; NAME/KEY: 5'UTR  
; LOCATION: (15)...(126)  
; FEATURE:  
; NAME/KEY: 3'UTR  
; LOCATION: (1468)...(1723)  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (154)...(213)  
; OTHER INFORMATION: function: membrane anchor (amino acids 10-29);  
; OTHER INFORMATION: product: hydrophobic amino acid stretch in GntI;  
; OTHER INFORMATION: standard_name: membrane anchor of a type II golgi  
; OTHER INFORMATION: protein.  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)-(14)  
; OTHER INFORMATION: function: use for cloning the cDNA library in  
; OTHER INFORMATION: Lambda ZAPII;  
; OTHER INFORMATION: product: EcoRI/NotI-cDNA adapter;  
; OTHER INFORMATION: number: 1.  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1724)...(1737)  
; OTHER INFORMATION: product: EcoRI/NotI-cDNA adapter;  
; OTHER INFORMATION: number: 2.  
US-10-615-144-3  
  
Query Match 100.0%; Score 1737; DB 8; Length 1737;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1737; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 GAATTCGGCGCGCCATTGACTTGTATCTTAAGTAAAGTAAATCCAGCGATGA 60  
DB 1 GAATTCGGCGCGCCATTGACTTGTATCTTAAGTAAAGTAAATCCAGCGATGA 60  
  
QY 61 AACACTATACTGAACACTGAGAGACTATTCGCTTTCTCTAAAGCCTTCAATCGAATT 120  
DB 61 AACACTATACTGAACACTGAGAGACTATTCGCTTTCTCTAAAGCCTTCAATCGAATT 120  
  
QY 121 CGCAGATGAGAGGGAACAAGTTTGTGCTGATTTCCGGTACCTCCTCATCTTCGGCTGCT 180  
DB 121 CGCAGATGAGAGGGAACAAGTTTGTGCTGATTTCCGGTACCTCCTCATCTTCGGCTGCT 180  
  
QY 181 GTCGCTTTCATCTACACACAGATCGGCTTTTTCGACACAGTCAAGTATGACAGATCGC 240  
DB 181 GTCGCTTTCATCTACACACAGATCGGCTTTTTCGACACAGTCAAGTATGACAGATCGC 240  
  
QY 241 CTTCGCTGCTGCAATTGAAGCAAAATCAATGTACAGCCAGACAGATTCCTTATTGAC 300  
DB 241 CTTCGCTGCTGCAATTGAAGCAAAATCAATGTACAGCCAGACAGATTCCTTATTGAC 300  
  
QY 301 CAGATTAGCTGCGACGAGGAGATAGTTGCTCTTGAAGAACAAATGAAGCGTCAGGAC 360  
DB 301 CAGATTAGCTGCGACGAGGAGATAGTTGCTCTTGAAGAACAAATGAAGCGTCAGGAC 360  
  
QY 361 CAGGAGTCCGACAAATTAAAGGCTCTTCTTCAGGATCTTGAAGTAAAGGCGATAAAAAG 420  
DB 361 CAGGAGTCCGACAAATTAAAGGCTCTTCTTCAGGATCTTGAAGTAAAGGCGATAAAAAG 420  
  
QY 421 TTGATCGGAAATGTACAGATGCCAGTGGCTGCTGTAGTTGTATGGCTTGCATTCGGGCT 480  
DB 421 TTGATCGGAAATGTACAGATGCCAGTGGCTGCTGTAGTTGTATGGCTTGCATTCGGGCT 480  
  
QY 481 GATTACCTGGAAAGACATTATAATCCATCTTAAATACCAATATCTGTTCGGTCAAAA 540
```

```
DB 481 GATTACCTGGAAAGACATTATAATCCATCTTAAATACCAATATCTGTTCGGTCAAAA 540  
QY 541 TATCCTCTTTTCATATCCAGGATGATCAATCTCTGATGTCAAGAGCTTGTCTTGAGC 600  
DB 541 TATCCTCTTTTCATATCCAGGATGATCAATCTCTGATGTCAAGAGCTTGTCTTGAGC 600  
QY 601 TATGATCAGCTGACGTATATGAGCACTTTGGATTTTGAACCTGTGCATCTGAAAGACCA 660  
DB 601 TATGATCAGCTGACGTATATGAGCACTTTGGATTTTGAACCTGTGCATCTGAAAGACCA 660  
QY 661 GGGGAGCTGATTCATACATACTACAAAATTGCACGTCAATTACAAGTGGCAATTCAGCTG 720  
DB 661 GGGGAGCTGATTCATACATACTACAAAATTGCACGTCAATTACAAGTGGCAATTCAGCTG 720  
QY 721 TTTTCAAGCATAAATTTTAGCCGTGTATCATCTAGAAAGATGATGGAATTTGCCCT 780  
DB 721 TTTTCAAGCATAAATTTTAGCCGTGTATCATCTAGAAAGATGATGGAATTTGCCCT 780  
QY 781 GATTTTTCGACTTTTTCGAGCTGAGCTACTCTTCTTGACAGACAAAGTCGATTATG 840  
DB 781 GATTTTTCGACTTTTTCGAGCTGAGCTACTCTTCTTGACAGACAAAGTCGATTATG 840  
QY 841 GCTATTTCTTCTTGGAAATGACAAATGCAGTTTGTCCAAAGATCCTTATGCTCTT 900  
DB 841 GCTATTTCTTCTTGGAAATGACAAATGCAGTTTGTCCAAAGATCCTTATGCTCTT 900  
QY 901 TACCGCTCAGATTTTTCGCGTCTTGGATGGATGCTTTTCAAAATCTACTCGGACGAA 960  
DB 901 TACCGCTCAGATTTTTCGCGTCTTGGATGGATGCTTTTCAAAATCTACTCGGACGAA 960  
QY 961 TTATCTCCAAAGTGGCCAAAGGCTTACTGGGACGCTGCTAGACTCAAGAGAAATCAC 1020  
DB 961 TTATCTCCAAAGTGGCCAAAGGCTTACTGGGACGCTGCTAGACTCAAGAGAAATCAC 1020  
QY 1021 AGAGGTGCAAAATTTATTCGCCAGAAAGTTTGAGAACATATATAATTTTGGTGAGCATGGT 1080  
DB 1021 AGAGGTGCAAAATTTATTCGCCAGAAAGTTTGAGAACATATATAATTTTGGTGAGCATGGT 1080  
QY 1081 TCTAGTTTGGGCGAGTTTTCAGAGAGTATCTTGGAGCAATTTAAACTAAATGATGTCAG 1140  
DB 1081 TCTAGTTTGGGCGAGTTTTCAGAGAGTATCTTGGAGCAATTTAAACTAAATGATGTCAG 1140  
QY 1141 GTTGATTGGAAGTCAATGACCTTACTTACCTTTTGGAGGACAAATTTACGTGAACACTTT 1200  
DB 1141 GTTGATTGGAAGTCAATGACCTTACTTACCTTTTGGAGGACAAATTTACGTGAACACTTT 1200  
QY 1201 GGTGACTTGGTTAAAAAGGCTAAGCCATCCATGGAGCTGATGCTGTCTTGAAGACATTT 1260  
DB 1201 GGTGACTTGGTTAAAAAGGCTAAGCCATCCATGGAGCTGATGCTGTCTTGAAGACATTT 1260  
QY 1261 AACATAGATGGTATGCGTATTCAGTACAGAGATCAACTAGACTTTGAAAATATCGCA 1320  
DB 1261 AACATAGATGGTATGCGTATTCAGTACAGAGATCAACTAGACTTTGAAAATATCGCA 1320  
QY 1321 CGGCAATTTGGCAATTTTGAAGATGGAAGGATGGTGTACCCAGTGCAGCATATAAGGA 1380  
DB 1321 CGGCAATTTGGCAATTTTGAAGATGGAAGGATGGTGTACCCAGTGCAGCATATAAGGA 1380  
QY 1381 ATAGTAGTTTTCGGTACCAAAAGCTCCAGACGTGATTTCCCTGTGGCCATGATTCGCTT 1440  
DB 1381 ATAGTAGTTTTCGGTACCAAAAGCTCCAGACGTGATTTCCCTGTGGCCATGATTCGCTT 1440  
QY 1441 CAACTCGGAATTGAAGATCTTAACTAAAGATATGATTCAGAGGAGCCGGGCAAAAT 1500  
DB 1441 CAACTCGGAATTGAAGATCTTAACTAAAGATATGATTCAGAGGAGCCGGGCAAAAT 1500  
QY 1501 TTTGACTTATTCGGTAGGATGCACTTAAACCATGATTTTACCAGTTACAT 1560  
DB 1501 TTTGACTTATTCGGTAGGATGCACTTAAACCATGATTTTACCAGTTACAT 1560  
QY 1561 ACAACGTTTAAATGTTATACGAGGAGCTCACCTGTTCTAGTGTGAAGGATATTCGGCTT 1620
```



```

Qy 786 TTTTGACTTTTTGAGCTGGAGCTACTCTCTCTTGACAGACAAAGTCGATTATGGCTAT 845
Db 712 TTTTGACTATTTTGAGGCTGGAGCTACTCTCTCTTGACAGACAAAGTCGATTATGGCTAT 771
Qy 846 TTCTTCTTGGAAATGACAATGGACAAAATGCAATTTGTGTCGAAGATCCTTATGCTCTTTACCG 905
Db 772 TTCTTCTTGGAAATGACAATGGACAAAAGGCAAGTTCGTCCAAGATCCTGATGCTCTTTACCG 831
Qy 906 CTGAGATTTTTCGCCGGTCTTGATGATGCTTTCAAAATCTACTTTGGGACGAAATATC 965
Db 832 CTGAGATTTTTCGCCGGTCTTGATGATGCTTTCAAAATCTACTTTGGGACGAAATATC 891
Qy 966 TCCAAAGTGGCCAAAGGCTTACTCGGAGCTGCTAAGACTCAAGAGAAATCACAGG 1025
Db 892 TCCAAAGTGGCCAAAGGCTTACTCGGAGCTGCTAAGGCTGAAGAAATCACAGG 951
Qy 1026 TCGACAAATTTATTCGCCCAGAAAGTTTGAGAACATATAATTTTGGTGAGCATGTTCTAG 1085
Db 952 TCGACAAATTTATTCGCCCAGAAAGTTTGAGAACATGACAAATTTTGGTGAGCATGTTCTAG 1011
Qy 1086 TTTGGGGCAGTTTTTCAAGCAGTATCTTTGAGCCAAATTAACATAATGATGTCAGGTTGA 1145
Db 1012 TTTGGGGCAGTTTTTAAAGCAGTATCTTTGAGCCAAATTAAGCTAAATGATGTCAGGTTGA 1071
Qy 1146 TTGGAAGTCAATGACCTTAGTTACCTTTTGGAGGACAATTAAGCTAAATGATGTCAGGTTGA 1205
Db 1072 TTGGAAGTCAATGACCTTAGTTACCTTTTGGAGGACAATTAAGCTAAATGATGTCAGGTTGA 1131
Qy 1206 CTTGGTTAAAAAGGCTTAAGCCCATCCATGGAGCTGATGCTGTTTGAAGCATTTAAACAT 1265
Db 1132 CTTGGTTAAAAAGGCTTAAGCCCATCCAGGAGCTGATGCTGTTTGAAGCATTTAAACAT 1191
Qy 1266 AGATGGTGATGTCGTTATTCAGTACAGAGATCAATAGACTTTGAAAATATCGCACGGCA 1325
Db 1192 AGATGGTGATGTCGTTATTCAGTACAGAGACCAACTAGACTTTGAAGATATCGCTCGACA 1251
Qy 1326 ATTTGGCATTTTCAAGATCGAAGGATGGTGATACCGTGACAGCATATAAGGAATAGT 1385
Db 1252 GTTTGGCATTTTGAAGATCGAAGGATGGTGATACCGTGACAGCATATAAGGAATAGT 1311
Qy 1386 AGTTTTCGGTACCAAGCTGCAGACGTGTATTCCTTGTGGCCATGATTCGCTTCAACA 1445
Db 1312 AGTTTTCGGTTCACACATCTAGACGTGTCTTCTGTTCCCTGATTCCTCTCGACA 1371
Qy 1446 ACTCGGAATGAAGATCTTAACAAAGATATGATTCAGGAGCCCGGCAAAATTTTGA 1505
Db 1372 ACTTGGAGTTGAAGATCTTAGCGAAGATATGATT--GGAGCCTGAGCAACAATTTAGA 1428
Qy 1506 CTTATTTGGTAGGATGATCGAGCTGACACTAAACCATGATTTT-----ACCAGTTAC 1558
Db 1429 CTTATTTGGTAGGATACATTTGAAAGAGCTGACACGAAAGATATGACTACAGTAGCTAC 1488
Qy 1559 ATACAAGCTTTTAATGTTATACGGAGGAGCTCACTGTTCTAGTGTGTAAGGGATATCGGC 1618
Db 1489 ATGCAACATTTTATGTTAATGGAAGGACCCACTGCT-----T 1527
Qy 1619 TTCTTAGTATTTGATGATATCAACAAACCTATTTTAAAGTGTTCGAAACATAAAG 1678
Db 1528 ATTGTGGAATGATGATATCATCACCACATCCTATTA--TTCAAAGTTTACAAACATAAAG 1585
Qy 1679 AGGAATGTAGCCCTGTAAAGACTA 1703
Db 1586 AGGAATGTTCGCCCTATATAAAACAA 1610

```

RESULT 3

US-10-615-144-5  
; Sequence 5, Application US/10615144  
; Publication No. US2004018127A1  
; GENERAL INFORMATION:  
; APPLICANT: Von Schaeuwen, Antje

; TITLE OF INVENTION: Plant GntI Sequences and the Use Thereof for the Production  
; of Plants Having Reduced or Lacking N-acetyl Glucosaminyl

```

; TITLE OF INVENTION: Transferase I(GntI) Activity
; FILE REFERENCE: 032266-003
; CURRENT APPLICATION NUMBER: US/10/615,144
; CURRENT FILING DATE: 2003-07-09
; PRIOR APPLICATION NUMBER: US/09/591,466
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: EP 98/08001
; PRIOR FILING DATE: 1998-09-12
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1854
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1185)...(1193)
; OTHER INFORMATION: function: Asn Codon is a potential glycosylation site;
; OTHER INFORMATION: product: Consensus sequence for N-glycosylation;
; OTHER INFORMATION: phenotype: N glycans modulate protein characteristics;
; OTHER INFORMATION: standard name: N glycosylation site;
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1185)...(1193)
; OTHER INFORMATION: label: pot-CHO;
; OTHER INFORMATION: note: absent in animal GntI sequences.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (135)...(1469)
; OTHER INFORMATION: codon_start: 135;
; OTHER INFORMATION: function: initiates complex N glycans on secretory glycoproteins;
; OTHER INFORMATION: EC_number: 2.4.1.101;
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (135)...(1469)
; OTHER INFORMATION: product: beta-1,2-N-acetyl glucosaminyl transferase I;
; OTHER INFORMATION: evidence: EXPERIMENTAL;
; OTHER INFORMATION: gene: cgl;
; OTHER INFORMATION: standard_name: gntI;
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (135)...(1469)
; OTHER INFORMATION: label: ORF;
; OTHER INFORMATION: note: first gntI sequence from Arabidopsis
; OTHER INFORMATION: (unpublished).
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (19)...(134)
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (1470)...(1848)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (157)...(215)
; OTHER INFORMATION: function: membrane anchor (amino acids 8-27);
; OTHER INFORMATION: product: hydrophobic amino-acid region in GntI;
; OTHER INFORMATION: standard_name: Membrane anchor of a type II Golgi protein;
; OTHER INFORMATION: note: identified by comparison with animal GntI sequences.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(18)
; OTHER INFORMATION: function: for preparation of a cDNA library in
; OTHER INFORMATION: Lambda ACT;
; OTHER INFORMATION: product: XhoI-cDNA-Adaptor;
; OTHER INFORMATION: number: 1.
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1849)...(1854)
; OTHER INFORMATION: product: XhoI-cDNA-Adaptor;
; OTHER INFORMATION: number: 2.
; US-10-615-144-5

```

Query Match 44.2%; Score 767.2; DB 8; Length 1854;

	Best Local Similarity	73.1%; Pred. No. 2.8e-216;	Mismatches 985; Conservative	0; Mismatches 363; Indels	0; Gaps	0;
Qy	139	AAGTTTTCGTGATTTCCGGTAGCCTTCCTCATCTTGCGCTGCCTGCCTTCATCTACACA	198			
Dd	141	AGGATCTGCTGTGACTTGAGATTCTTCTCATCCGGCAGCTTTCATGTTTCATCTACATC	200			
Qy	199	CAGATCGGCTTTTTCCGACACAGTCAGAAATATGCAAGTCGCCCTGCTGCTCCAATTGAA	258			
Dd	201	CAGATGAGGCTTTTCAGAGCGCAATCACAGTAGTCAGATCGCCTCAGTTCCGCTATCGAA	260			
Qy	259	GCGAAAAATCATGTACAAAGCCAGACACAGATTGCTTTATTGACCAGATTAGCTGCGAGCAA	318			
Dd	261	TCTGAGAACCATTTGCACTAGTCAAAATGCGAGGCTCATAGATGAAGTTAGCATCAAAAG	320			
Qy	319	GGAAGAATAGTTGCTCTTGAAGAACAAAATGAAGCGTCAAGACACGAGAGTCCCAGCAATTA	378			
Dd	321	TCGCGGNTTGTGCGCTCGNAGATATGAAGAACCGCCAGGACGAAGAACTTGTGCAGCTT	380			
Qy	379	AGGCTCTTGTTCAGGATCTTGAAGCTTAAGGGCATATAAAGTTGATTCGGAATGTACAG	438			
Dd	381	AAGGATCTAATCCAGACGCTTTGAAAAAAAAGGAATAGCAAAAATCTCACTCAAGGTGACAG	440			
Qy	439	ATGCCAGTGCCTGCTGTAGTTGTTATGGCTTGCAATCGGGCTGATTACCTTGGAAAAGACT	498			
Dd	441	ATGCCGTGCGCTGCTGTAGTGGTTATGGCTGCGAGTCGTGCAGACTATCTTGAAGGACT	500			
Qy	499	ATTAAATCCATCTTAAATACCAAAATATCTGTTCGCGTCAAAAATATCCTCTTTTCATATCC	558			
Dd	501	GTTAAATCAGTTTAAATATCAAACTCCCGTTCCTTCAAAATATCCTCTATTATATCT	560			
Qy	559	CAGGATGGATCAATCTCTGATGTCAGGAAGCTTGTCTTTTGAGCTATGATCAGCTGACGTAT	618			
Dd	561	CAGGATGGATCTGATCAAGCTGTCAGAGCAAGTCAATTGAGCTATAATCAATTAACATAT	620			
Qy	619	ATGCAGACCTTGGATTTTGAACCTGTGCATCTGAAAGACCAGGGGAGCTGATGTCATAC	678			
Dd	621	ATGCAGACCTTGGATTTTGAACCAAGTGGTCACTGAAAGGCTCGGTGAACCTGACTCGGTAC	680			
Qy	679	TACAAAAATTCGACGTCAATTAAGTGGGCATTTGATCAGCTGTTTTCACAGCAATAATTTT	738			
Dd	681	TACAAGATTGCACTGCTACTCAAGTGGGCACCTGGAACGATGTTTTCACAAACAATAATTT	740			
Qy	739	AGCCGTGTTATCAPACTAGAAAGTATATGGAATTCGCCCTGATTTTTTTTGACTTTTTTT	798			
Dd	741	AGTCGAGTGATTAFACTAGAAGACGATATGGAATTTGCTCCAGACTTCTTTGATTACTTTT	800			
Qy	799	GAGGCTGGAGCTACTCTTTTGACAGACAGACGTCGATTAATGCGTATTTCTTCTTGGAT	858			
Dd	801	GAGGCTGCAGCTAGTCTCATGGATAGGGATAAAACCAATTATGCGTCTTCATCATGGAAT	860			
Qy	859	GACAAATGSACAAATGCAAGTTTGTCCAAGATCTTTATGCTCTTTTACCGCTCAGATTTTTTT	918			
Dd	861	GATAAATGGACAGAAGCAGTTTGTGCATGATCCCTATGCGCTATACCGATCAGATTTTTTT	920			
Qy	919	CCCCGTCTTTGGATGGATGCTTTCAAAATCTACTTTGGGACGAAATATATCTCAAAAGTGSCCA	978			
Dd	921	CCTGGCTTTGGTGGATGCTCAAGAGATCGACTTTGGGATGAGTTATCACCAAAGTGSCCA	980			
Qy	979	AAGGCTTATCTGGGACGACTGCTTAAGACTCAAAAGAAATCAAGAGTCCAGCAATTTATT	1038			
Dd	981	AAGGCTTATCTGGGATGATTGGCTGAGACTAAAGGAAAAACCAATAAGGCCGCCCAATTCATT	1040			
Qy	1039	CGCCAGAGAGTTTGCAGACATATAATTTTGGTTCAGCATGGTCTAGTTTGGGGCAGTTT	1098			
Dd	1041	GCAACGGAAAGTCTGTAGAACATACAAATTTTGGTGAACATGGGTCTAGTTTGGGACAGTTT	1100			
Qy	1099	TTCAAGCAGTATCTTTGAGCCAAATTAACATAAATGATGTCACAGGTTGATTTGGAAGTCAATG	1158			
Dd	1101	TTCAAGTCAATCTTTGGAACCTTAATAAGCTAAACGATGTGACGTTGACTTGGAAAGCAAG	1160			
Qy	1159	GACCTTAGTTACCTTTTGGAGGACAAATTAACGTGAAAACACTTTTGGTGTGACTTGGTTAAAAAG	1218			

Db 343 AAGACCTGCAGAGCTGATTGATTAAGTGCAGGTTCCGGTGGCTGCTGTTGTGATCATGG 402  
Qy 467 CTTGCAATCGGGCTGATTAACCTGAAAGACTATAAATCCATCTTAAATACCAAAATAT 526  
Db 403 CATGTAATCGTCTGATTAACCTGGAGAGGACTATAATCTGTATTGAAATACCAAGGC 462  
Qy 527 CTGTTGGCTCAAAATATCCCTTTTTCATATCCAGGATGGATCACATCCTGATGCAGGA 586  
Db 463 CCATTTCTTCAAGATATCCCTTTTATTTGATCTCAGGATGGATCAATCCAAATGTTTAA 522  
Qy 587 AGCTTGTCTTGGCTATGATCAGCTGACGTATATGACGACCTTGGATTTTGAACCTGTGC 646  
Db 523 GTAAGGCTTGGCTATGATCAGTATCTTATATGACGACCTTGGATTTGAACCACTTC 582  
Qy 647 ATACTGAAGACCAAGGGAGCTGATGTCATATCAAAAATTCACGCTCATTTACAAGTGG 706  
Db 583 AAACCTGAACGGCTGGAGAGTTAACTGTTACTACAAAATTCACGCTCATTTACAAGTGG 642  
Qy 707 CATTGGATCAGCTCTTTTACAGCATAAATTTAGCCGTGTTATCATACTAGAGATGATA 766  
Db 643 CTCTAGATCAACTGTTCTACAAGCATAACTTCAGCCGGTGTGATCAITCTTGAAGATGACA 702  
Qy 767 TGGAAATGCCCCCTGATTTTGTGACTTTTGTGAGGCTGGAGCTACTCTCTTGACAGAG 826  
Db 703 TGGAAATAGCACCTGATTTCTTGATTTTGTGAGCTGACGACCTCTCTTGACAGAG 762  
Qy 827 ACAAGTCGATTTATGGCTATTTCTTCTTGGAAATGCAATGGACAAATCAGTTTCCAA 886  
Db 763 ATAAATCAATATGCTGTTTCTCTCATGGAATGCAATGGACAAAGCAGTTTGTACATG 822  
Qy 887 ATCCTTATGCTTTTACCGCTCAGATTTTTCGGGTCTTGGATGATGCTTTCAAAT 946  
Db 823 ATCCATATGAATTTATGCTCAGACTTCTTCTGGAATAGGATGGATGCTGCCAGAT 882  
Qy 947 CTACTTGGAGCAATTAATCTCAAAGTGGCCAAAGGCTTACTGGGACGACTGGCTAAGAC 1006  
Db 883 CTACATGGATGACTATCACCAAAATGGCCAAAG-----CTGGATGACTGGTTGAGAC 938  
Qy 1007 TCAAAGAGATCAAGAGGTGCAAAATTTATTCGCCCAGAAAGTTTGGAGCAATATAATT 1066  
Db 939 TTAAGAGATCAAAAGGAGCAGATTTATCCGGCCGGAAGTATGCAGAACATATAATT 998  
Qy 1067 TTGCTGAGCATGTTCTAGTTTGGGCGAGTTTTCAGACGATATCTTGGCCATTTAAAC 1126  
Db 999 TTGCTGAGCATGTTCTAGTTTGGGACAGTTTTTTAAAGCAATTTCTTGGCCAAATCAAG 1058  
Qy 1127 TAAATGATGTC----- 1137  
Db 1059 TGAATGATGCAAGGTACACCTGCTTTATGAAACAATACAGTCTTAAATATGTTCTTAT 1118  
Qy 1138 -----CAGTTGATGGAAAGTCAATGAACC 1162  
Db 1119 TATTAGAATGATGCTATTTCATTTGGATTTTGGAAAGTTGATTGGAAATCAATGATT 1178  
Qy 1163 TTAGTTACCTTTTGGAGACAAATTAAGTGAACACTTTGGTACTGTTGTTTAAAGGCTA 1222  
Db 1179 TGAGCTATTACTGGAGGATAAATATCTATGCACTTTGGCAAGCTTGTAAAGAAAGCTA 1238  
Qy 1223 AGCCCATCCATGGAGCTGATGCTGCTTGAAGCAATTTAAACATAGATGGTATGCGGTA 1282  
Db 1239 CACTGTCTATGGAGCTGACATGTTCTAAAGGCAATATAATATAGATGGGATGCGGTA 1298  
Qy 1283 TTCAGTACAGAGATCAACTAGACTTTGAAATATTCGACCGCAATTTGGCATTTTGAAG 1342  
Db 1299 TCAAATATGAAGATCAGTCAGACTTTGAAACATTTGCTCGCAATTTGGTATATTTCAAG 1358  
Qy 1343 AATGGAGGATGGTGTACCACTGCAGCATATAAGNATAGTATTTCCGGTACCAAA 1402  
Db 1359 AGTGGAGGATGGTGTGCCAGGACAGCCTATAAGGAGTAGTCTGTTTCAGATATCAAA 1418  
Qy 1403 CGTCAGACGCTGATTTCTCTGTGGCCATGATTCGCTTCAACAACTCGGAATTTGAAGATA 1462

Db 1419 CCTCAGCAGTATATTCTTGTGTTGGTCCAGAAATATTTGAAGCTACTTCAGATCGAAGAGT 1478  
Qy 1463 CTTAA 1467  
Db 1479 CTTAA 1483  
RESULT 5  
US-10-437-963-21004  
; Sequence 21004, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 21004  
; LENGTH: 1594  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_26317C.1  
US-10-437-963-21004

Query Match 38.5%; Score 668.2; DB 7; Length 1594;  
Best Local Similarity 70.0%; Pred. No. 6.7e-187;  
Matches 914; Conservative 0; Mismatches 388; Indels 3; Gaps 1;

Qy 147 CTGTGATTTCCGTACTCTCTCATCTTGGTCTGTGCTGCTGCCTTATCTACACAGATGCG 206  
Db 15 CTGCGACCTCCGATCCTCTCTGTCGCGCGCTGCAGCCTTCTATCTACATCCAAAGTCCG 74  
Qy 207 GCTTTTGGACACAGTCAGAAATATGCAGATCCGCTTGTGCTGCAATTCGAAGCAGAAAA 266  
Db 75 CTTTTTTTCTACTCAGTCCCACTACGCCGCCCTCGCACAGCAGAAAAATCCGAAAA 134  
Qy 267 TCATTGTACAAGCCAGACAGATTTGCTTTATTTGACAGATTAGCTTCGAGCAAGGAAGAT 326  
Db 135 CCAATGCACCAGTCAACTCCGCTCTTATCGATCAAGTCAGCTCTCAGCAGGAGAAGAT 194  
Qy 327 AGTTGCTCTTGAAGAACAAATGAAGCGTCAAGACAGGAGTGGGCAAAATTAAGGGCTCT 386  
Db 195 TGTGCGGTTGGAAAGAGATGAAGATACGCCAAGACGAAGAACGTTGTACACCTCAAGATTTT 254  
Qy 387 TGTTCAGGATCTTGAAGTAAGGGCATAAAAAGTTGATCGGAAATGTACAGATGCCAGT 446  
Db 255 AATCCAGATCTCGAAAAAAGGAGCGTGCAGACTTATAGTAAACAACATGTGGCTCTGT 314  
Qy 447 GGCTGCTAGTGTGTTATGGCTTGCATCGGGCTGATTAACCTGGAAAAAGACTATTAATC 506  
Db 315 TGCTGCTGTTGTTGTAATGCTTGCATCGACCAGACTATTTGCAAGGACAGTTGAATC 374  
Qy 507 CATCTTAAATACCAAAATATCTGTTGGTCAAAAATATCTCTTTTCTATATCCAGGATGG 566  
Db 375 TATCCTGAAGTATCAGACATCAGTGGCTTCAAAGTTTCCACTATTATATCTCAGGATGG 434  
Qy 567 ATCAGATCCTGATGTCAGGAAGCTGCTTTGAGCTATGATCAGCTGACGTATATGACGA 626  
Db 435 AATTAATGGGAAGTGAAGAAAAAGCTTTGAGTTTATTAATGAATATACATATATGACGA 494  
Qy 627 CTTTGGATTTTGAACCTGTGCATATCTGAAAGACCGAGGAGCTGAATTCATATCAAAAAT 686  
Db 495 TTTAGATCTTGAACCTGTGCGCACTGAAAGGCCAGGAGAGTTGATAGCATATTACAAGAT 554

QY 687 TGCAGCTCATTTACAGTGGGATTTGGATCAGCTGTTTACAGCATATTTTACGCCGT 746  
DB 555 AGCCAAACATTTACAAATGGGCTTGGATGAGTATTTATCAACACACAACTTTGGCGGAGT 614  
QY 747 TATCATATCTAGAGATGATATGGAATTTGCCCTGATTTTGTGACTTTTGTGAGGCTGG 806  
DB 615 AATAATTTCTGGAGATGATATGGAGATCGCCCTGATTTTGTGACTACTTTCGAGGCTGC 674  
QY 807 AGCTACTCTTTGTGACAGAGCAAGTATGCTATTTGCTTTTGTGGAATGCAATGG 866  
DB 675 AGCTAAATTTCTTTGATAACAGTAAGACAAATCATGGCCGTTTCTTCTGGAATGCAATGG 734  
QY 867 ACATATGAGTTTGTCCAGATCTTTATGCTCTTTTACCGCTCAGATTTTTCCTCCGCTCT 926  
DB 735 ACAAAAGCAATTTGTGTAGCCCAAAAGCTCTTTTACCGCTCGGATTTTCTTCTGGAAT 794  
QY 927 TGGATGGATGCTTTTCAAAATCTACTTTGGAGCAATTTATCTCAAAAGTGGCCAAAGGCTTA 986  
DB 795 TGGATGGATGTTAAACAGCTTACATGGATGGAGCTGTCACTAAGTGGCCCAAGCTTA 854  
QY 987 CTGGACAGCTGGCTTAAGACTCAAAAGAAATCAGAGAGTGCAGAAATTTATTCGCCCA 1046  
DB 855 TTGGGATGACTGGGTGAGGCTAAAGAGAGGTACACAGAGACCGGCAATTTATTCGCCCGA 914  
QY 1047 AGTTTGCAGACATATATTTTGGTGGAGCATGTTTCTAGTTTGGGGCAGTTTTCAGCA 1106  
DB 915 AGTCTGCAGAACATACAAATTTTGGGCAACATGGGTCAAGCATGGGGCAATTTCTTCAGACA 974  
QY 1107 GTATCTTGAGCCAAATTAATCTAAATGATGTCAGGTTGATTTGGAAGTCAATGGACCTTAG 1166  
DB 975 GTACTTGGNACCAATCAAGCTTAATGATGCCCATATCAAGTGAATTTCTGAGGACCTGAG 1034  
QY 1167 TTACTTTTGGAGGCAATTAAGTGAACACTTTTGGTGTGCTTTGTTTAAAGGCTAAGCC 1226  
DB 1035 CTACCTCAAGGAGGCAAGTCTCTGATCCAAATTTGGGAAGACGTCGCTAGTGGCCACCCC 1094  
QY 1227 CATCATGAGCTGATGCTGCTTGAAGCAATTTAAGCATGATGATGCTGCTATTCA 1286  
DB 1095 TCTCCATGATCTGATGCGCGCTTGAAGCCCAATATGGAATGGAATGGAATGGAATGCA 1154  
QY 1287 GTACAGAGATCAATAGACTTTGAAATATCGACGGCAATTTGGCAATTTTGAAGAATG 1346  
DB 1155 GTATACGATCAGGAGACTTTGAGCGGATAGCTCGTCAATTTGGAATATTTGAAGAATG 1214  
QY 1347 GAAGGATGCTGACCAAGTGCAGCATATTAAGAAATAGTATGTTTCCCGGTACCA---AAC 1403  
DB 1215 GAAGGATGCAATTTCCAAAGAACAGCTTATAAGGAGTAGTAGTCTTCCCGGTACAAGAGTAG 1274  
QY 1404 GTCCAGAGCTGATTTCTTGTGGCCATGATTCGCTTCAACAACT 1448  
DB 1275 TCGAAGACGAATATACCTCGTTGGTCCGATTTCCCTCAGTCAGCT 1319

## RESULT 6

US-10-425-114-487  
; Sequence 487, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 487  
; LENGTH: 1882

## ; TYPE: DNA

; ORGANISM: Zea mays

## ; FEATURE:

; OTHER INFORMATION: Clone ID: 700077454\_F11

US-10-425-114-487

Query Match 38.1%; Score 661.8; DB 7; Length 1882;

Best Local Similarity 68.4%; Pred. No. 5.9e-185;

Matches 931; Conservative 0; Mismatches 427; Indels 3; Gaps 1;

QY 147 CTGTGATTTCCGGTACCTCTCATCTTGGCTGCTGCTCGCCCTTCACTACACACAGATCGG 206  
DB 160 CTGGACCTCCGCT 219  
QY 207 GCTTTTTCGACACAGCTCAGAAATATGAGATCGGCTTGTCTGCTGCTCAATTTGAAGCAGAAA 266  
DB 220 CCTCTTGTCTACTAGTCCCATATGCGGACCGTCTTGGGAAGCAGAAAGATCTGAATA 279  
QY 267 TCATTTGACAAAGCCAGACAGATTTGCTTATTTGACAGATTTAGCCTTGCAGCAAGGAGAT 326  
DB 280 TCAATGCACCACTCAATTTAAAGTCTCTTATCGATCAAGTCAGCATGCGACGAGGAAGAT 339  
QY 327 AGTTGCTCTTGAAGAACAAATGAAGCGTCAAGACCGAGGTCGCGACAAATTTAAGGGCTCT 386  
DB 340 TGTAGCATTTGAAGAGATTAAGGTACGCAAGATGAAGACGTGCGACATCTGAGGATTTT 399  
QY 387 TGTTCAGGATCTTGAAGTAAGGGCATAAAAAGTTGATCGGAAATGTACAGATGCCAGT 446  
DB 400 GATTAAGGATCTTGAANAAGAGGTGTGCAAGACTACTAGACAAGATGTGTTCCCGT 459  
QY 447 GGCTGCTGTAGTGTATTTGGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 506  
DB 460 TGCTGCTGTGCTATATATGCTTGAATGCTTGAATGCTTGAATGCTTGAATGCTTGAATGCT 519  
QY 507 CATCTTAAATACCAATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 566  
DB 520 TATCTGAAGTATCAGACATCAGTGTCTTGAAGTCTTGAAGTCTTGAAGTCTTGAAGTCT 579  
QY 567 ATCATCTCTGATGTGCAAGACTTGTCTTGAAGTCTGATCAGCTGATGATGATGATGATGATG 626  
DB 580 AGCAATGGAGCTGTGAAGATTAAGCTTTGGAATATAACAAATAACATACATCAGCA 639  
QY 627 CTGGAATTTGAACCTGTGCAATCTGAAAGACCGAGGAGCTGATGCTGCTGCTGCTGCTGCTG 686  
DB 640 TGTGGATCTTGAACCTGTGCAACCTGAAAGGCGCAGGAGATTTGACAGCATATTTACAAGAT 699  
QY 687 TGCAGCTCATTTACAAGTGGGCAATTTGGATCAGCTGTTTACAAGCATATTTTAGCCGTGT 746  
DB 700 TGTCTAAACATTTATAAGTGGGCTTGGCAACACCTATTTCAATTAACATAATTTTGTCTGAGT 759  
QY 747 TATCATATCTAGAAGATGATATGGAATTTGCCCTGATTTTGTGACTTTTGTGAGGCTGG 806  
DB 760 AATCAATCTAGAAGATGATGAGATTTGCCCGGATTTTTCGACTACTTTCGAGGCTGC 819  
QY 807 AGCTACTCTTCTTGACAGAGCAAGTCAATTAAGCTATTTCTTCTTGAATGCAATGG 866  
DB 820 AGCTAAACTACTTGTATGATGACAGCAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 879  
QY 867 ACAATGAGATTTGTGCAAGATCTTATGCTCTTTACCGCTCAGATTTTTCCTCCGCTCT 926  
DB 880 GCAGGAAGCAGTTTGTAAACGCCCAAAAGCTCTTTTACCGCTCAGATTTTCTTCTCCGCTCT 939  
QY \*927 TGGATGGATGCTTTCAAAATCTACTTTGGAGCAATTTATCTCAAAAGTGGCCAAAGGCTTA 986  
DB 940 TGGATGGATGTTTAAACAAATGCTTGAATTTGTCACCAAAAGTGGCTTAAGCTTA 999  
QY 987 CTGGACGACTGGCTTAAGACTCAAAAGAGAAATCAGAGGTGCAATTTTATTCGCCCA 1046  
DB 1000 TTGGGATGATTTGGTGGAGCTTAAGAGGATGATGATGATGATGATGATGATGATGATGATG 1059  
QY 1047 AGTTTGCAGAACATATAATTTTGGTGGAGCATGCTGCTTGTGTTGGGCGAGTCTTTTCAAGCA 1106  
DB 1060 AATTTGTAGAACATACAAATTTTGGCAAGCATGATGATGATGATGATGATGATGATGATGATG 1119







```
Db 1525 TGAGAGGTAGATTTTACCTTGGATGGCAATATCATG 1565
|||||
RESULT 8
US-10-767-795-2499
; Sequence 2499, Application US/10767795
; Publication No. US20040181830A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Cao, Yongwei
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53234)B
; CURRENT APPLICATION NUMBER: US/10/767,795
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 117596
; SEQ ID NO 2499
; LENGTH: 1211
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: GOSHI-09MAY01-C89123_1
US-10-767-795-2499

Query Match 38.1%; Score 661; DB 8; Length 1211;
Best Local Similarity 75.1%; Pred. No. 7.8e-185;
Matches 851; Conservative 0; Mismatches 280; Indels 2; Gaps 2;

QY 127 ATGAGAGGGAACAAGTTTGTGTCATTTCCGGTACCTCCTCATCTTGCTGCTCGCC 186
DB 80 ATAAAATGGCTAAGCTAAGCTGCGATTTCCGGTACCTTCATCCCGCTGCATCGTA 139

QY 187 TTCATCTACACACAGATCGGCTTTTTCGACACAGTCAGATATGCAAGATCGCCTTGT 246
DB 140 TTCATATACATCCAGATCGGCTTTTCCACAGCAATCTGAATATGCAAGATCGAATGGCA 199

QY 247 GCTGCAATTGAAGCAAAATCATTTGTACACGACAGCAGATTCCTTATTCACAGATT 306
DB 200 GAAGCGGTGCTGACAGACCAATTTGTACGAGTCAATTCGACTACTAATTTGATCAAAAT 259

QY 307 AGCTGCGACAGGAAGATAGTGTCTTTGAAGAACAAATCAAGCGTCAGACACAGGAG 366
DB 260 AGTATGCAACAGAGCAAAATCGTGGCCCTTGAAGAGGGAAGAGCAAGGATCAGGAG 319

QY 367 TGCCGACAAATTAAGGCTCTTGTTCAGGATCTTGAAGTAAGGCGATAAAAAAGTTGATC 426
DB 320 TGCGCACATTTGAAGACTCTTGTTAATGATCTTGAAGAAAGACTCTTCAAAGAGTAAT 379

QY 427 GGAATGTACAGATGCGAGTGTCTGTAGTTGTATGCTTGCATCGGCTGATTAC 486
DB 380 GACAAAACACAGGTACTCTGCGAGCTGTAGTAATCATGCGTGCATCTGCTGATTAT 439

QY 487 CTGGAAGACACTATTAATCCATCTTAAATACCAATATCTTGTGCGTCAAAATATCT 546
DB 440 CTGGAAGACAGCTTCGCTGTCTTAAAGTATCAGAGCTCTGTGCTTCGAAGTATCCA 499

QY 547 CTTTTTCATATCCAGGATGGATCACATCTCTGTATGTCAGGAAGCTTGTCTTTGAGCTATGAT 606
DB 500 CTTTGTGTATCTCAGGATGATCAGATCCAAAGGTTTAAACTAAGGCTTTGAGTTATAAG 559

QY 607 CAGTCAAGCTATATGAGCACTTGGATTTTGAACCTGTGCTATCTGAAAGACACAGGGAG 666
DB 560 GAGCTAACTTATATGAGCACATAGATATGATTCCTGCTGCATACAGACCGGCTGGGAA 619

QY 667 CTGATTCATCTACAAAATTCACCTCATTAAGTGGGCTTGGATCAGCTGTTTTAC 726
DB 620 TTGATCGCATCTACAAAGATTTGCCGCTCACTACAAATGGGCTATGGATGATTTCTAC 679

QY 727 AAGCAAAATTTAGCCGCTGTATCATCTAGAGATGATATGAAATTCGCCCTGATTTT 786
DB 680 AAGCAAAATTTGACCGAGTAAATACCTTGAAGATGATATGAAATTTGCCCTGATTTT 739
```

```
QY 787 TTTGACTTTTTTTGAGGCTGGAGCTACTCTTTTGACAGACAGTGCATATATGGCTATT 846
DB 740 TTTGATTACTTTGAGGCGAGCTGCTGCCCTTCTCGACAAGGACAGTCAAATATATGGCTGTT 799

QY 847 TCTTCTTTGGAATGACAAATGCAAAATGCTTTGCTCAAGATCTTTATGCTCTTTACCGC 906
DB 800 TCCTCATGGAATGCAATGCGCAAAAGCAGTTTGTGTATGACCCATATGCACTTTATCGC 859

QY 907 TCAGATTTTTTTCCCGGTCTTTGGATGATGCTTTCAAAATCTACTTGGGACGAATTTATCT 966
DB 860 TCAGATTTCTTTCTGCTTTGGCTGGATGCTTAACTGTAATCTGTATGGAATGAGCTATCA 919

QY 967 CCAAAGTGGCCAAAGGCTTACTGGGACGCTGGCTTAAGACTCAAAAGAGAAATCAGAGGT 1026
DB 920 CCAAATGGCCAAAGCTTACTGGGATGACTGGTTGAGATTTAAAAAATAATCACAATGGT 979

QY 1027 CGACAATTTATTCGCCAGAAAGTTTGCAGACATATATTTTGGTGAGCAGTGTCTTAGT 1086
DB 980 CGACAATTTCTTCTGCTCTGGAATATGCAACAATATTTTGGTGAGCATGGTTTCAAGC 1039

QY 1087 TTGGGCGAGTTTTTCAAGCAGTATCTTGAGCCAATTTAAACTTAAATGATGTCCAGTTGAT 1146
DB 1040 ATGGGCGAGTTTTTGGAAAAATACCTTGCACCTATTAAAGATGAATGAGCT-CAGGTGGAC 1098

QY 1147 TGGAGTCAATGAGACCTTTAGTTACCTTTTGGAGACAAATTAAGTGAACACACTTTGGTGAC 1206
DB 1099 TGGAGTCCA-GGATTTGAGCTACCTAACTGAGGAAAAATATGCCCAATACTTTGCAGAC 1157

QY 1207 TTGGTTAAAAAGGCTAAGCCCATCATGGAGCTGATGCTGTCTTTGAAAGCAAT 1259
DB 1158 ATTTTGAAGCGCAAAACCTGTCCCTGGAAACAGATTTCTCTCTTATGGCATT 1210
```

```
RESULT 9
US-10-424-599-93080
; Sequence 93080, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 93080
; LENGTH: 2293
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_55065C.1
US-10-424-599-93080
```

```
Query Match 35.4%; Score 614.6; DB 7; Length 2293;
Best Local Similarity 71.8%; Pred. No. 7.1e-171;
Matches 821; Conservative 0; Mismatches 314; Indels 8; Gaps 1;

QY 333 TCTTCAAGAACAAATGAAGCGTTCAGGACGAGGTCGCGACAAATTTAAGGGCTCTTTGTTC 392
DB 620 TCTTTCAGAGAGCAGAAACGCTCGAGACCAAGAAATGTGGACAAACAAGTCTCTTGTACA 679

QY 393 GGATCTTGAAGTAAGGGCATAAAAAGTTGATCGGAAATGTACAGATGCCAGTGGCTGC 452
DB 680 AGATCTTGAAGAAAGAACCTTCAGAGGCTGATTTGATAAAGTGCAGGTTCCGGTGGCTGC 739

QY 453 TGTAGTTGTTATGGCTTGCATTCGGGCTGATTTACTCTGGAAGAGACTATTAATCCATCTT 512
DB 740 TGTGTGATCATGGCATGTAATCGTCTGATTAACCTGGAGAGGACTATTAAATTTCTGTCAT 799
```

```
Qy 513 AABATACCAATATCTGTTGGTCAAAATATCCTCTTTTCATATCCAGGATGATCACA 572
Db |||
Qy 800 AACTTCAGCCGTGTAATCATCTTGAAGATGACATTTGAAATAGCAACGGATTTCTTCAAT 859
Db |||
Qy 573 TCCTGATGTCAGGAAGCTTGTCTTGGATATGATCAGCTGACGTATATGACGACCTTGGGA 632
Db |||
Qy 860 TATTGTAAGCTGCAGCGAGTCTTCTTGAGATGATAAATCCATTATGCGTGTCTTCTT 919
Db |||
Qy 633 TTTTGAACC-----TGTGCATACCTGAAAGACAGGGGAGCTGATGCAATACAAA 684
Db |||
Qy 920 ATGGAATGCCACTTGGATTTTGAACCTGAACGGCTGGAGAGTTAACTGCTTACTACAAA 979
Db |||
Qy 685 ATTGCAGCTCAATACAGTGGCATTTGATCAGCTGTTTACAGACATAATTTTAGCCGT 744
Db |||
Qy 980 ATTGCAGCTCAATACAGTGGCTCTAGATCAACTGTTTACAAAGCATAACTTCAGCCGT 1039
Db |||
Qy 745 GTTATCATCTAGGAAGATGATGGAATTCGCCCTGATTTTGTGACTTTTGTGAGGCT 804
Db |||
Qy 1040 GTATCATCTTGAAGATGACATGGAATAGCACTGATTTCTTGTATTTTGAAGCT 1099
Db |||
Qy 805 GGAGCTACTCTTGTGACAGACAGCAAGTCGATTATGCTATTTCTTGTGAATGCAAT 864
Db |||
Qy 1100 GCAGCGACTCTCTTGAAGGATAAATCCATTATGCTGTTTCTCATGGAATGACAA 1159
Db |||
Qy 865 GGCAAAATGCAATTTGTCAGATCCTTATGCTCTTTACCGCTCAGATTTTTCGCCGT 924
Db |||
Qy 1160 GGCAAAAGCAGTTGTACATGATGTCATGAACTTTATGCTCTGACTCTTCTTCTCGGA 1219
Db |||
Qy 925 CTTGGATGGATGCTTTCAAAATCTACTTGGGACAAATTTATCTCCAAAGTGCCAAAGGCT 984
Db |||
Qy 1220 TTAGATGGATGCTGGCCAGATCTACATGGATGAGCTATCACCAAAATGCCAAAGCT 1279
Db |||
Qy 985 TACTGGGACGCTGGCTAAGACTCAAGAGAAATCACAGAGTCGACAAATTTATTCGCCCA 1044
Db |||
Qy 1280 TACTGGGATGCTGGTTGAGACTTAAAGAGAAATCACAAAGGACGACAGTTTATCCGCC 1339
Db |||
Qy 1045 GAAGTTTGCAGAACATATAATTTTGGTGAGCATGGTCTAGTTTGGGCGAGTTTTCAG 1104
Db |||
Qy 1340 GAAGTATGCAGAACATATAATTTTGGTGAGCATGGTCTAGTTTGGGACAGTTTTCAG 1399
Db |||
Qy 1105 CAGTATCTTGAGCCAAATTAATGATGTCCAGGTTGATTGGAAATCAATGCACTT 1164
Db |||
Qy 1400 CAATTTCTTGAGCCAAATCAAGCTGATGATGTCAAGTTGATTGGAATTAATGGATCTG 1459
Db |||
Qy 1165 AGTTACCTTTTGGAGGACAAATTAAGTGAACACATTTTGGTGACTGGTTTAAAGGCTAAG 1224
Db |||
Qy 1460 AGCTATTTACTGGAGGATAAATATTTCTATGCACCTTTGCGAACGTTGTTAAGAAAGCTACA 1519
Db |||
Qy 1225 CCCATCCATGAGCTGATGCTGCTTGAAGCATTTTAACATAGATGGTGATGTCGCTATT 1284
Db |||
Qy 1520 CCTGTCTATGGAGCTGACATGGTCTTAAGGCATCTAATATAGATGGCGATGTGCGCATC 1579
Db |||
Qy 1285 CAGTACAGAGATCAACTAGACTTTGAAATATCGCAGGCAATTTGGCATTTTGAAGAA 1344
Db |||
Qy 1580 AAATATAAGATCAGTCAAGCTTTGAAACAACTTCTCACCAGTTTGGTATATTTCAAGAG 1639
Db |||
Qy 1345 TGAAGGATGTTGATCCAGTGCAGCATATTAAGGAAATAGTATGTTTCCGGTACCAACG 1404
Db |||
Qy 1640 TGAAGGATGTTGTTGCCCGGACAGCATATAAAGGAGTAGTTCGTTTTCAGATATCAAAAC 1699
Db |||
Qy 1405 TCCAGAGCTGATTTCTTGTGGCCATGATTTGCTTCAACAACTCGGAATTTGAAGATCT 1464
Db |||
Qy 1700 TCAGACGATATATCTTGTGGTCCAGAACTTTTGAAGCTACTTCAGATCGAAGAGTCT 1759
Db |||
Qy 1465 TAA 1467
Db 1760 TAA 1762
```

```
; GENERAL INFORMATION:  
; APPLICANT: Kovalic, David K.  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND OTHER MOLECULES ASSOCIATED WITH  
; FILE REFERENCE: 38-21(53377)B  
; CURRENT APPLICATION NUMBER: US/10/739,930  
; CURRENT FILING DATE: 2003-12-18  
; NUMBER OF SEQ ID NOS: 11088  
; SEQ ID NO 5511  
; LENGTH: 1898  
; TYPE: DNA  
; ORGANISM: Triticum aestivum  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(1898)  
; OTHER INFORMATION: unsure at all n locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: TRIAE-23APR03-CLUSTER98_776  
US-10-739-930-5511
```

Query Match 33.5%; Score 581.2; DB 8; Length 1898;  
Best Local Similarity 66.4%; Pred. No. 5.3e-161;  
Matches 887; Conservative 0; Mismatches 429; Indels 19; Gaps 3;

```
Qy 147 CTGTGATTTCCGGTACCTCCTCATCTTGGCTGCTGCGCTTCATCTACACAGATGCG 206
Db |||
Qy 1778 CTGCGACATCCGGCTCCTCCTGTCGCCGCCGCGCTTTCATCTACATCCAAGTGG 1719
Db |||
Qy 207 GCTTTTGGGACACAGCTCAGAAATATGCAGATCGCTTGTCTGTGCAATTTGAAGC 260
Db |||
Qy 1718 GCTGTTCTCCACCAGTCTTCTCTTCTTCTTAATCCCACTTTCAGACGCCGAGC 1659
Db |||
Qy 261 -----AGAAATCATTTGTACAAGCCAGACCAAGTTGCTTATTTGACAGATTAGCCT 311
Db |||
Qy 1658 CTTGCGGGAAGACAGAGAAATGCGAGAGCCAGCTGCACGCCATGANGCAGCCAGGTCAGCG 1599
Db |||
Qy 312 GCAGCAAGGAAGAAATAGTTGCTCTTCAAGAACAAATGAAGCGTCAGAGCCAGGAGTGGCG 371
Db |||
Qy 1598 CCAGCAGGAGAAATCGCCGCACTTGAAGAGATGAAGGTACGCCAAGACGAGGAGGCTGT 1539
Db |||
Qy 372 ACAATTAAGGGCTCTTGTTCAGGATCTTGAAAGTAAGGGCATAAAAAGTTGATCGGAA 431
Db |||
Qy 1538 GCAGCTCAAGATTTTAATCCAGATCTTGAAGAGAGAGTTTACAGACCTTAAACAACA 1479
Db |||
Qy 432 TGTACAGATGCCAGTGGCTGTGTATGTTATGCTTGCATTCGGGCTGATTAACCTGGA 491
Db |||
Qy 1478 GAACGTGGTTCCTGTCGCGCTGTTGTCATAATGCTTGCATTCGACACCAAGCTATTGCA 1419
Db |||
Qy 492 AAAGACTATTAATCCATCTTAAATACCAATATCTGTTGCTGCAAAATATCCTCTTTT 551
Db |||
Qy 1418 GAGGACAGTGGAAATCTATCCTCAAGTACCAAGAACGATGCTTCAAGATTTCCACTATT 1359
Db |||
Qy 552 CATATCCCAGGATGGATCACATCCTGATGTACAGAAAGCTTGTCTTGAGCTTATGATCAGCT 611
Db |||
Qy 1358 TATATCACAGGATGGAACAAATGGAGAAGTAAAGAAAGAGCTTGTGATTACATCAAAT 1299
Db |||
Qy 612 GAGCTATATGACACA-CTTGGATTTTGAACCTGTGTGATATCTGAAAGACCAAGGGGAGCTGA 670
Db |||
Qy 1298 AACATTTATGACAGATGTGGATCTTTGAGCCTGTGCGCATGAAAGACCAAGAGAAACG 1239
Db |||
Qy 671 TTGCATCTACAAATTTGACGCTCATTAAGTGGGCATTTGGATGAGCTGTTTTTACAGC 730
Db |||
Qy 1238 TTGCATATTACAGATAGCTAACCACTATAAAATGGGCTTTGGATGAGCTATTATTAAAGC 1179
Db |||
Qy 731 ATAATTTTAGCCGTGTTATCATACTAGAAGATGATATGGAATTTGCCCTGATTTTTTTTG 790
Db |||
Qy 1178 ATGATTTTGTGCGAGTAATCATTTCTGGAAGATGACATGGAGATCGCCCGAGATTCTTTG 1119
Db |||
Qy 791 ACTTTTTTAGGCTGAGCTACTCTTTCTTGACAGACAGCAAGTCGATTAAGGCTATTCTTT 850
Db |||
Qy 1118 ACTACTTTGAGGCTGCAGCGGAAATTAATTGACACTGACAAAGACAATAAATGGCTGTTTCAT 1059
Db |||
Qy 851 CTTGGAATGACAAATGACAAATGCAAGTTTGTCCCAAGATCCTTATGCTCTTTTACCGCTCAG 910
Db |||
```

```
Db 1058 CTTGGATGATATGGGCAAAAGCAGTTCGTTTATGACCCAAAGCTCTTTACCGTTCGG 999
Qy 911 ATTTTTCCTCCGGTCTTGGATGAGTCTTTCAAAATCTACTTGGGAGCAATATCTCCAA 970
Db 998 ACTTCTTTCGGGGCTTGGATGAGTCTTAAAGGATCAACATGATGGAGCTGTCCCAA 939
Qy 971 AGTGCCAAAGGCTTACTGGGACGACTGGCTAAGACTCAAGAGAAATCACAGAGGTGCAC 1030
Db 938 AGTGCCCAAGGCTTATTTGGGATGACTGGGTGAGGCTAAAGAGAGGTACACAGAGATCGG 879
Qy 1031 AATTATTTCCGCCAGAGTTTTCGAAACATATAATTTTGGTGAGCATGGTTCTAGTTTGG 1090
Db 878 AGTTTATTCGCCAGAGATGATGAGGACATACAACTTTTGGCGAGCATGGATCAAGCATGG 819
Qy 1091 GCGAGTTTTCAAGCAGTATCTTGAGCCAAATTAACATAATTAATGATGTCAGGTTGATGGA 1150
Db 818 GACAATCTTTGATCAATACCTTGAACCTATCAAGTTAAATGATGCTCATATTGACTGGA 759
Qy 1151 AGTCAATGAGCTTAGTTACCTTTTGGAGGACAAATACGTGAACACCTTTTGGTGAATTGG 1210
Db 758 ATTCGAGGACTGAGCTACCTCAAGGAGGACAACTTTTGAACCAATTCGGGAAAGACG 699
Qy 1211 TTAAGAGGCTAAGCCCATCCATGAGCTGATGCTCTTGAAGACATTTAAACATAGATG 1270
Db 698 TGGCTAGCGCCACACCTGTGATGATCGATGCTCTTGAAGGCCACACATCTGGATG 639
Qy 1271 GTGATGTCGGTATTCAGTACAGAGATCAACTAGACTTTGAAATATCGCAGGCAATTTG 1330
Db 638 TGGAGCTAAGGATTCAGTATGACATCAGGCGGACTTTGAGCGTATAGCTCGGCAATTTG 579
Qy 1331 GCATTTTGAAGATGGAAGTGTGTACACGTCGAGCATATAAAGGAATAGTAGTTT 1390
Db 578 GAATATTTGAAGATGGAAGGACGCTGTTCACGGGGCGCTTACAAAGCGTGTGGTGT 519
Qy 1391 TCCGGTACCA---AACGTCGACGCTGATTCCTTTTGGCCCATGATTCGCTTCAACAC 1447
Db 518 TCCGGTACAAAGGCGAGCGGAGACGATATACCTGGTGGTCCGACTCTTTCGCCAAC 459
Qy 1448 TCGGAATGAAGATA 1462
Db 458 TCGGGGTTAGCTTA 444
```

## RESULT 11

```
US-10-424-599-93082
; Sequence 93082, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 93082
; LENGTH: 1414
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(1414)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_55067C.1
US-10-424-599-93082
```

Query Match 33.2%; Score 577; DB 7; Length 1414;  
Best Local Similarity 72.3%; Pred. No. 7.8e-160;

```
Matches 814; Conservative 0; Mismatches 261; Indels 51; Gaps 3;
Qy 139 AAGTTTTCGCTGATTTCCGCTACCTCTCATCTTGGCTGCTGCTGCTCATCACACA 198
Db 333 AAGTTTTCGCTGATTTCCGCTTCTCTCTGATTCAGCTGGGCTTCTCATCTATATC 392
Qy 199 CAGATGCGGCTTTTGGGACACAGTTCAGAAATATGAGATCGCTTGTCTGCTGCAATGAA 258
Db 393 CAGATGAGGCTTTTGGCAACACAAATCAGAAATATGCTGATCGCTCGCTGAGCTATCGAA 452
Qy 259 GCAGAAATCATTTGTAACAGCCAGACAGATGCTTATTCAGCAGATTTAGCTCTCAGCAA 318
Db 453 GCTGAAACCATTTGTAACAGTCAACGCGATCGCTGATTCAGATAGCTTGCACAA 512
Qy 319 GGAAGAATAGTTGCTCTTTGAAGA-----ACAAATGAAGCGTCAGGACAGGAGTGCCG 371
Db 513 GGACGAATTTGGCCCTTGGAGAGAAACGTTAGAGCAGAAACGTCGAGACCAAGAAATGTGG 572
Qy 372 ACAATTAAGGCTCTTGTTCAGGATCTTGAAGATGTAAGGCAATAAAAGTTGATCGGAAA 431
Db 573 ACAACAAAGTCTCTTGTACAAGATCTTGAAGAAAGAACCTGCGAGAGGCTGATTGATAA 632
Qy 432 TGTAACAGATGCCAGTGGCTGCTGTAGTTGTATGGCTTGAATCGGCTGATTAACCTTGA 491
Db 633 AGTCAGGTTCCGGTGGCTGCTGTGTGATCATGGCATGTAATCGTCTGATTAACCTTGA 692
Qy 492 AAAGACTATTAAATCCATCTTAAATACCAAAATATCTGTTCGCTCAAAATATCTCTTTT 551
Db 693 GAGGACTATTAAATCTGTATTAATTAACAAAGGCCCATTTCTTCAAGATATCTTTATT 752
Qy 552 CATATCCAGGATGGATCAATCTCTGATGTGAGAAAGCTGCTTTGAGCTATGATCAGCT 611
Db 753 TGTATCTCAGGATGGATCAAACTCGAAACTGGAGATAAACTGAGGGAGC-----801
Qy 612 GACGTATATGCGACACTTGGATTTTGAACCTGTCGACTACTGAAGACCCAGGGAGCTGAT 671
Db 802 -----AAATTTGCAATTTTGCCTAAAGAA-----TAT 828
Qy 672 TGCATCTACAAAATTCGACGTCATTAACAAGTGGGCTATGGATCAGCTGTTTAAACAAGA 731
Db 829 TCAATGATTCGCTTATCAGGTCAATTAACAAGTGGGCTCTAGATCAACTGTTCTACAAGCA 888
Qy 732 TAATTTTAGCGCTGTATCATATCTAGAAGATGATATGGAATATGCCCCCTGATTTTGA 791
Db 889 TAACCTTCAGCGCTGTATCATCTTGAAGATGATATGGAATATGCACTGATTTCTTGA 948
Qy 792 CTTTTCGAGCTGGAGTACTCTTCTTGAAGAGACAGTTCGATTAATGGCTATTTCTTC 851
Db 949 TTAATTTGAAGCTGCAGCGAGTCTCTTTGAGAAAGATAAATCCATTTATGGCTGTTTCTC 1008
Qy 852 TTGGAATGACAATGACAAATGCAATGCTTTTGTCCAAAGATCCTTATGCTCTTTACCGCTCAGA 911
Db 1009 ATGGAATGACAATGACAAAGAGGTTTGTATGATGATCATATGAACTTTATCGCTCTGA 1068
Qy 912 TTTTTCCTCCGCTCTTGGATGGATGCTTTCAAAATCTACTTTGGGACGAATATCTCCAAA 971
Db 1069 CTTCTTCTCTGATTTAGGATGGATGCTGCGCAGATCTACATGGAATGAGCTATCAACAAA 1128
Qy 972 GTGGCCAAAGGCTTACTGGGACGACTGGCTAAGACTCAAGAGATCAAGAGGTCGACA 1031
Db 1129 ATGGCTAAAGCTTACTGGGATGACTGGTGTGAGACTTAAAGAGAAATCAAAAGGACGACA 1188
Qy 1032 ATTTATTCGCCAGAGTTTCAGAAACATATAATTTTGTGAGCATGTTCTTAGTTTGGG 1091
Db 1189 GTTTATTCGCCGCCGAAAGTATGCAAAACATATAATTTTGTGAGCATGTTCTTAGTTTGGG 1248
Qy 1092 GCAGTTTTTCAAGCAGTATCTTGAAGCAATTTAAATCTAAATGATGTCAGGTTGATGAAA 1151
Db 1249 ACAGTTTTTAAAGCAATATCTTGAAGCAATCAAGCTGAATGATGATCAAGGTTGATGAAA 1308
Qy 1152 GTCAATGAGCCTTAGTTACCTTTTGGAGGACAAATTAACGTGAAGAACTTTGGTGAAGTGT 1211
Db 1309 ATTAATGGATCTGAGCTATTTACTGGAGGATAAAATTAATCTATGCACTTNTGCAATGTTAT 1368
```

Qy 1212 TAAAGGCTAAGCCCATCCATGAGCTGATGCTGCTTGAAGCA 1257  
Db 1369 TAAGAAAGCTACACCTGTCTATGGAGCTGACATGGTTCTAAAGCA 1414

RESULT 12  
US-10-424-599-93078  
; Sequence 93078, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 93078  
; LENGTH: 954  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_55063C.1  
US-10-424-599-93078

Query Match 30.4%; Score 527.8; DB 7; Length 954;  
Best Local Similarity 78.2%; Pred. No. 2.5e-145; Indels 0; Gaps 0;  
Matches 634; Conservative 0; Mismatches 177;

Qy 139 AAGTTTCTGTGATTTCCGGTACCTCCTCATCTTGGCTGCTGCGCTTCATCTACACA 198  
Db 144 AAGTTTCTGTGATTTCCGGTCTCTTCTTGGTGGCGTGGCGTTCATCTATATC 203

Qy 199 CAGATGGCGCTTTTGGGACACAGTCAGAAATATGAGATCGCCCTGCTGCTGCAATGAA 258  
Db 204 CAGATGAGCGCTTTTCGCAACTCAATCAGAAATATGCTGATCGCTAGCTGAGCTATTGAA 263

Qy 259 GCAGAAATCAATGTACAGCCAGACAGATGCTTATTGACACAGATTAGCCTGCACAA 318  
Db 264 GCTGAAACCAATGTACAAAGTCAACGCGATCACTGATTGATCAGATTAGCTTGCAACA 323

Qy 319 GGAAGAATAGTTGCTCTTGAAGACAAATGAAGCGTGCAGACAGGAGTGCCGCAATTA 378  
Db 324 GGACGAATTTGGCCCTAGAGAAGAGCGGAAACGCCGAGACCAAGATGTGGACAAATG 383

Qy 379 AGGCTCTTTGTCAGGATCTTGAAGTAAGGGCATAAAAAGTTGATCGGAAATGTACAG 438  
Db 384 AAGTCTCTGTGACAGATCTTGAAGAAAGAACCTACAGAGGCTGATTGATAAATGCAG 443

Qy 439 ATGCCAGTGGCTGCTGTAGTTGTTATGCTTGCATCGGCTGATTACCTGGAAAGACT 498  
Db 444 GTTCGGTGGCTGCTGTGATCATGGCATGTAAATCGTGTGATTACCTAGAGAGACT 503

Qy 499 ATTAATCCATCTTAAATATACCAATATCTGTTCGGTCAAAATATCTCTTTTCATATCC 558  
Db 504 ATTAATCTGTATTAAATATACCAAGGCCCATTTCTCAAGATATCTTTATTGTATCT 563

Qy 559 CAGATGATCACATCTCTGATGTGAGGAGCTTGTCTTTGAGCTATGATCAGCTGACGTAT 618  
Db 564 CAGATGATCAAAATCCAAATGTAAAGTAAAGGCTTTGAGCTGTGATCAGTTATCTTAT 623

Qy 619 ATGCAGCACTTGGATTTGAACCTGTGCATACGAAAGCCAGGGAGCTGATTGTCATAC 678  
Db 624 ATGCAGCACTTGGATTTGAACCAAGTCAAACTGAACGGCTGGAGAGTTAACTGCTTAC 683

Qy 679 TACAAAATGACCGCTCATTTACAAGTGGGCATTGGATCAGCTGTGTTTTTCAAGCATAAATTT 738  
Db 684 TACAAAATGACCGCTCATTTACAAGTGGGCTCTAGATCAACTGTTCTACAGCATAACTTC 743

Qy 739 AGCCGTGTTATCATACTAGAAAGATGATAGAAATTTGCCCTGATTTTGTGACTTTTTT 798  
Db 744 AGCCGTGTTATCATTTCTTGAAGATGACATGGAATAAGCACCTTGATTTCTTTGATTA 803

Qy 799 GAGGCTGGAGCTACTCTTCTTGACAGAGACAAAGTCGATTATGCTGATTTCTTCTTGAAT 858  
Db 804 GAAGCTGCAGGACTCTCTTGCACAGGATAAATCCATTATGGCTGTTTCTCTCATGGAAT 863

Qy 859 GACAATGGACAAATGCAAGTTTGTCCAAGATCCTTATGCTCTTTTACCGCTCAGATTTTTT 918  
Db 864 GACAATGGACAAAGCAGTTTGTACATGATCCATATGAACTTTATCGCTCAGACTTCTTT 923

Qy 919 CCGGTCTTTGGATGGATGCTTTCAAAATCTA 949  
Db 924 CCTGGATTAGGATGGATGCTGGCCAGATCTA 954

RESULT 13  
US-10-437-963-21003  
; Sequence 21003, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 21003  
; LENGTH: 1849  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_26316C.1  
US-10-437-963-21003

Query Match 19.8%; Score 343.6; DB 7; Length 1849;  
Best Local Similarity 56.6%; Pred. No. 2.1e-90;  
Matches 780; Conservative 0; Mismatches 529; Indels 69; Gaps 5;

Qy 147 CTGTGATTTCCGTACTCTCTCATCTTGGCTGCTGCTGCTTTCATCTACACAGATGCG 206  
Db 418 CTGCGACCTCCGGATCCTCTCGTCCGCCGCTGCAGGCTTTCATCTACATCCAAGTCCG 477

Qy 207 GCTTTTGGACACAGCTCAGAAATATGCAGATCGCTTGTGCTGCAATTTGAAGCAGAAAA 266  
Db 478 CCTTTTCTACTCAGTCCCACTACGCCGACCGCTCGCACAGCAGAAAAATCCGAAAA 537

Qy 267 TCATTGTACAAGCAGACCCAGATTGTTATTGTACAGATTAGCTTCAGCAGAGAGAAT 326  
Db 538 CCAATGCACCAGTCAACTCGCTCTTGTATCGATCAAGTCAGCTCTCAGCAGGAGAAGAT 597

Qy 327 AGTTGCTCTTGAAGACAAATGAGCGTCAGGACAGGAGTCCGCAATTAAGGCTCT 386  
Db 598 TGTGCGGTTTGGAAAGAGATGAAGATACGCCAAGACGAAGAACGTTGTACACCTCAAGATTTT 657

Qy 387 TGTTCAGGATCTTGAAGTAAAGGCAATAAAAAATTTGATCGGAAATGTACAGATGCCAGT 446  
Db 658 AATCCAGATCTCGAAAAAAGGAGCGTGCAGACCTTAGTAAACAACTGTGCTCTGT 717

Qy 447 GGCTGCTGTAGTTGTTATGCTTGCATCGGGCTGATTAACCTGGAAAAAGACTATTAAATC 506  
Db 718 TGCTGCTGTGTTGTAATGCTTGCATCGACAGACTATTTTGCAAAGGACAGTTGAATC 777

Qy 507 CATCTTAAATACCAATATATCTGTTGGTCAAAATATCTCTTTTTCATATATCCAGGATGG 566

Db 778 TATCTGAGGATACAGCATCAGTGGCTTCAAAGTTTCCATATTTATATCTCAGGATGG 837  
Qy 567 ATCATCTCTGATGTCAGGAAGCTTGTCTTGTAGCTATGATCAGCTGACGTATATGCGACA 626  
Db 838 AATAAATGGAGAGTGAAGAAAAGCTTGTAGTTATTAATGAATAACATATATGCGAGAT 897  
Qy 627 CTTGGATTTGAACTGTGCATACGAAGACCCAGGAGCTGATGTCATACATAAAT 686  
Db 898 CCAATGAGTCATATCAGGAAATCTAGGTCCAATGATATATAAATGAGCAGTGAGAGCA 957  
Qy 687 TGCAGCTCAT--TACAAGTGGGCAATGGATCAGCTGTGTTTACAAGCATAAATTTAGCCGT 744  
Db 958 AGGCCATTAATCATACCCAACTTGTGATGTATTTTGAAGGGAAGCTTGTATGCCAC 1017  
Qy 745 GTATCATCATAGAGATGATGAAATGTGCCCTGTATTTTGTGACTTTTGTGAGG-- 802  
Db 1018 TGTACTGTACTTCCCTCCACATAACATAGGCAATGAGCAGTCCGATATTTGTAAGA 1077  
Qy 803 ----CTGGAGCTACTCTTGTGACAGACAACTCGATTATGCTATTTCTTCTGGAATG 859  
Db 1078 ACATGGGAAATATTTTACGTAAAGACTCTCAGTGAAATTAACAATGGGCCCTTGGATGA 1137  
Qy 860 ACAATGGA--CAATGCGAGTTTGTCCAAGATCCTTATGCTCTTTTACCGCTCAGATTTTTT 918  
Db 1138 GCTATTCATCAACACAACTTGTGGCGAGTATATAATCTGGAAGATGATATGAGATCGC 1197  
Qy 919 CCGGCTTTGAGATGATCTTCAAAATCTACTTGGGAGCAATTAATCTCCAAAGTGGCA 978  
Db 1198 CCTGTATTTTGTACTACTTGGAGGCTGAGCTAAATTTACTTGAACGATAAGACAAT 1257  
Qy 979 AAG----- 982  
Db 1258 CATGGCCGTTTCTTCTGGAATGACAAATGGAACAAAGCAATTTGTTAGACCCAACTT 1317  
Qy 983 --CTTACTGGGAGCTGCTGAAGACTCAAAGAGAAATCAAGAGGTGCAAAATTTATTGG 1040  
Db 1318 TACTTATTTGGATGATCTGGTGAGCTAAAGAGGTACACAGAGCCGCAATTTATTGG 1377  
Qy 1041 CCCAGAGTTTGCAGAACATATAATTTTGGTGAGCATGTTCTAGTTTGGGCGAGTTTTT 1100  
Db 1378 CCGGAAAGTCTGAGAACATACAAATTTTGGCGAAACATGGGTCAAGCATGGGGCAATCTT 1437  
Qy 1101 CAAGAGTATCTTGAGCCAAATTAACATAATGATGCTCAGGTTGATGGAAGTCAATGCA 1160  
Db 1438 CAGACAGTATTTGGAAACCAATCAAGCTAAATGATGCCCATATCAAGTGGAAATTTCTGAG 1497  
Qy 1161 CCTTAGTACCTTTTGGAGGACAAATACGTGAACAACTTTGGTGACTTGGTTAAAAAGGC 1220  
Db 1498 CTTGAGCTACCTCAAGAGGACAAAGTTCTGTATCCAAATTTGGGAAGAGCTCGCTAGTGC 1557  
Qy 1221 TAAGCCCATCCATGGAGCTGATGCTCTTGTGAAGCAATTAACATAGATGGTATGTCG 1280  
Db 1558 CACCCCTCTCCATGGATCTGATGCCGCTTGAAGGCCCAATATGATGCGGATGTAAG 1617  
Qy 1281 TATTCAGTACAGAGATCAACTAGACTTTGAAATATCGCAGGCAATTTGGCAATTTTGA 1340  
Db 1618 GATCCAGTATACGATCAGGAAGACTTTGAGCGGATAGTCTCGCAATTTGGAATATTTGA 1677  
Qy 1341 AGAATGGAAGGATGTTTACCGCTGAGCATATAAAGGAATAGTATGTTTCCGGTACCA 1400  
Db 1678 AGAATGGAAGGATGGCAATTTCCAAGACAGCTTATAAAGGATAGTATGTTCCGGTACA 1737  
Qy 1401 ---AAGCTCAGAGCTGATTTCTTGTGGCCATGATTCGCTTCAACAACTCGGAAT 1455  
Db 1738 GAGTAGTCGAGAGCAATATACCTCGTTGGTCCGATTTCCCTCAGTCAGCTTAGGTTT 1795

RESULT 14  
US-10-767-701-5996  
; Sequence 5996, Application US/10767701  
; Publication No. US20040172684A1  
; GENERAL INFORMATION:

; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof For Plant Improvement  
; FILE REFERENCE: 38-21 (53535)B  
; CURRENT APPLICATION NUMBER: US/10767,701  
; CURRENT FILING DATE: 2004-01-29  
; NUMBER OF SEQ ID NOS: 63128  
; SEQ ID NO 5996  
; LENGTH: 769  
; TYPE: DNA  
; ORGANISM: Sorghum bicolor  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)-(769)  
; OTHER INFORMATION: unsure at all n locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-CLUS30971\_1  
; US-10-767-701-5996  
  
Query Match 12.8%; Score 222; DB 7; Length 769;  
Best Local Similarity 64.7%; Pred. No. 1.6e-54;  
Matches 330; Conservative 0; Mismatches 180; Indels 0; Gaps 0;  
  
Qy 147 CTGTGATTTCCGGTACCTCTCATCTTGGCTGTCTGCGCTTCATCTACACAGATCGG 206  
Db 259 CTGGACCTCGCTCTCTCTCTGCTGCGCTTCATCTACATCCAAAGGCG 318  
Qy 207 GCTTTTTCGACACAGTCAATATGAGATGCGCTTCTGCTGCAATTTGAGCAAAA 266  
Db 319 GGTCTCTGCTGCTCAGTCCCATTTATGAGATCGTCTTCAGAAAGCAGAAAGATCTGAAA 378  
Qy 267 TCATTGTACAAGCCAGACGAGATTTGTTTATGACAGATTTAGCTTCGACAGGAAGAAT 326  
Db 379 TCAATGACACCTGCTCACTAAAGTCTCTGATCGATCAAGTCAAGTCAAGGAGAGAT 438  
Qy 327 AGTTGCTCTTGAAGAACAAATGAAGCGTCAGGACCAAGAGTCCGACAAATTAAGGCTCT 386  
Db 439 TGTAGCATTTGGAAGATTAAGGTACGCAAGATCAAGAACGCGGCACATCTGAAGATTT 498  
Qy 387 TGTTCAGATCTTGAAGTAAGGGCATATAAAGTTGATCGGAATGTACAGATCCAGT 446  
Db 499 GATAAAGGATCTTGAAGAAAGAGGTGTCAGAAAGTACTAGACAAATAATGTGGTTCTGT 558  
Qy 447 GGCTGCTGTAGTTGTTATGGCTTGAATCGGCTGATACCTGGAAGAGACTTATAAATC 506  
Db 559 TGCTGCTGTGTCAATAATGGCTTGAATCGGCACTATTTAGAGAGAACAGTTGAATC 618  
Qy 507 CATCTTAAATAACCAATATCTGTTGCGTCAAAATATCTCTTTTATATCCAGGATGG 566  
Db 619 TATCTGAAATATCAGACACAGTTGCTTCAAGTTTCGCTTTTATATCTCAGATGG 678  
Qy 567 ATCATCTCTGATGTCAGGAAGCTTGTGAGCTATGATGAGCTGACGTATATGCGACA 626  
Db 679 AGCAATGGAGCTGTCAGAAAGGAAAGCTTTGGAATATAAACAATAATATATGCGACA 738  
Qy 627 CTTGATTTTGAACCTGTGCTACTGTAAG 656  
Db 739 TGTGATCTTGAACCTGTGCAAACTGAAAG 768  
  
RESULT 15  
US-09-923-876-5569  
; Sequence 5569, Application US/09923876  
; Patent No. US20020013958A1  
; GENERAL INFORMATION:  
; APPLICANT: Lalgudi, Raghunath V.  
; APPLICANT: Kamigaki, Laura Y. (Ito)  
; APPLICANT: Sherman, Bradley K.  
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN SEEDLING  
; FILE REFERENCE: PL-0012-1 CON  
; CURRENT APPLICATION NUMBER: US/09/923,876

```
; CURRENT FILING DATE: 2001-08-06
; PRIOR APPLICATION NUMBER: 09/298,329
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/085,331
; PRIOR FILING DATE: 1998-05-05
; NUMBER OF SEQ ID NOS: 6332
; SOFTWARE: PERL Program
; SEQ ID NO 5569
; LENGTH: 278
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20020013958A1 700457259H1
US-09-923-876-5569
```

```
Query Match      10.5%; Score 182; DB 3; Length 278;
Best Local Similarity 78.4%; Pred. No. 6.5e-43;
Matches 218; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

Qy 625 CACTTGGATTGTAACCTGTGCATCTGAAAGACGAGGGAGCTGATTGCATACATAAAA 684
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1 CATGTGGATCTTGAACCTGTGAAACTGAAAGGCCGAGAGAATTGACAGCATATTACAAG 60

Qy 685 ATTGCACGTCAATTACNAGTGGGCATTGGATCAGCTGTTTACAGCATAAATTTAGCCGT 744
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
61 ATTGCTAAACATTATAAGTGGGCCCTTGGACAACCTATTTCATTAAACATAATTTTGCTCGA 120

Qy 745 GTTATCATCTAGAAGATGATGGAATTCGCCCTGATTTTGTGACTTTTGTGAGGCT 804
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
121 GTATCATCTTAGAAGATGACATGGAGATTGCCCGGATTTTTCGACTACTTCGAGGCT 180

Qy 805 GGAGCTACTCTTCTTGACAGAGACAAGTCGATTATGGCTATTCTTCTTGGAAATGACAAT 864
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
181 GCAGCTAAACTACTTGTATAATGACAAGACGATTATGGCTGTTTCGCTCTTGGAAATGACAAT 240

Qy 865 GGACAAATGCAGTTTGTCCAGATCCCTTATGCTCTTTA 902
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
241 GGGCAGAGAGCAGTTTGTTTACGACCCCAAAAGCTCTTTA 278
```

Search completed: December 14, 2005, 13:30:43  
Job time : 1313 secs

November 2005

Published\_Applications\_Nucleic\_Acid and Published\_Applications\_Amino\_Acid database searches now generate two sets of results each. The Published\_Applications databases have been split into two parts to reduce the amount of time required for their daily updates. This results in more machine time being available for processing searches.

Newly published applications will appear in the Published\_Applications\_New databases; older published applications make up the Published\_Applications\_Main databases.

Searches run against Nucleic Acid Published\_Applications produce two sets of results, with the extensions **.rnpbm** (Published\_Applications\_NA\_Main) and **.rnpbn** (Published\_Applications\_NA\_New).

Searches run against Amino Acid Published\_Applications produce two sets of results, with the extensions **.rapbm** (Published\_Applications\_AA\_Main) and **.rapbn** (Published\_Applications\_AA\_New).

Page blank (uspio)